DOCUMENT RESUME

ED 084 657 88 EA 005 636

AUTHOR Pyecha, John N.
TITLE Minnesota Educational Assessment: A Comprehensive Planning Study.
INSTITUTION Research Triangle Inst., Durham, N.C. Center for Educational Research and Evaluation.
SPONS AGENCY Minnesota State Dept. of Education, St. Paul. Div. of Planning and Development.
PUB DATE Jan 73
NOTE 108p.

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS *Academic Achievement; Data Analysis; Data Collection; Data Processing; *Educational Assessment; Educational Planning; Educational Programs; Evaluation; Evaluation Methods; *Evaluation Techniques; Information Dissemination; Information Utilization; Management; *Program Effectiveness; State Departments of Education; *State Programs
IDENTIFIERS Elementary Secondary Education Act Title III; ESEA Title III; Minnesota; NAEP; National Assessment of Education Progress

ABSTRACT

The proposed program represents a vehicle whereby Minnesota can conduct a comprehensive assessment of the State's educational progress. It can provide a means of periodically monitoring achievement in the cognitive, affective, and psychomotor domains. The program is modeled after the National Assessment of Education Progress (NAEP), an ongoing educational project designed to give educators and the lay public a better look at those knowledges and skills that American youth have acquired. NAEP provides for a systematic, continuous, census-like survey of knowledges, skills, understandings, and attitudes as exhibited by students and young adults in four age levels and across ten different subject areas. By following the NAEP model, Minnesota can hopefully reduce the Minnesota student assessment results to NAEP results for students in the nation as a whole, as well as for those in the Central Region; and take advantage of exercise administration, data collection, sampling, and data analysis methodologies developed by NAEP. (Portions of pages 29 and 86 may reproduce poorly.) (Author/WM)

FILMED FROM BEST AVAILABLE COPY

ED 084657

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

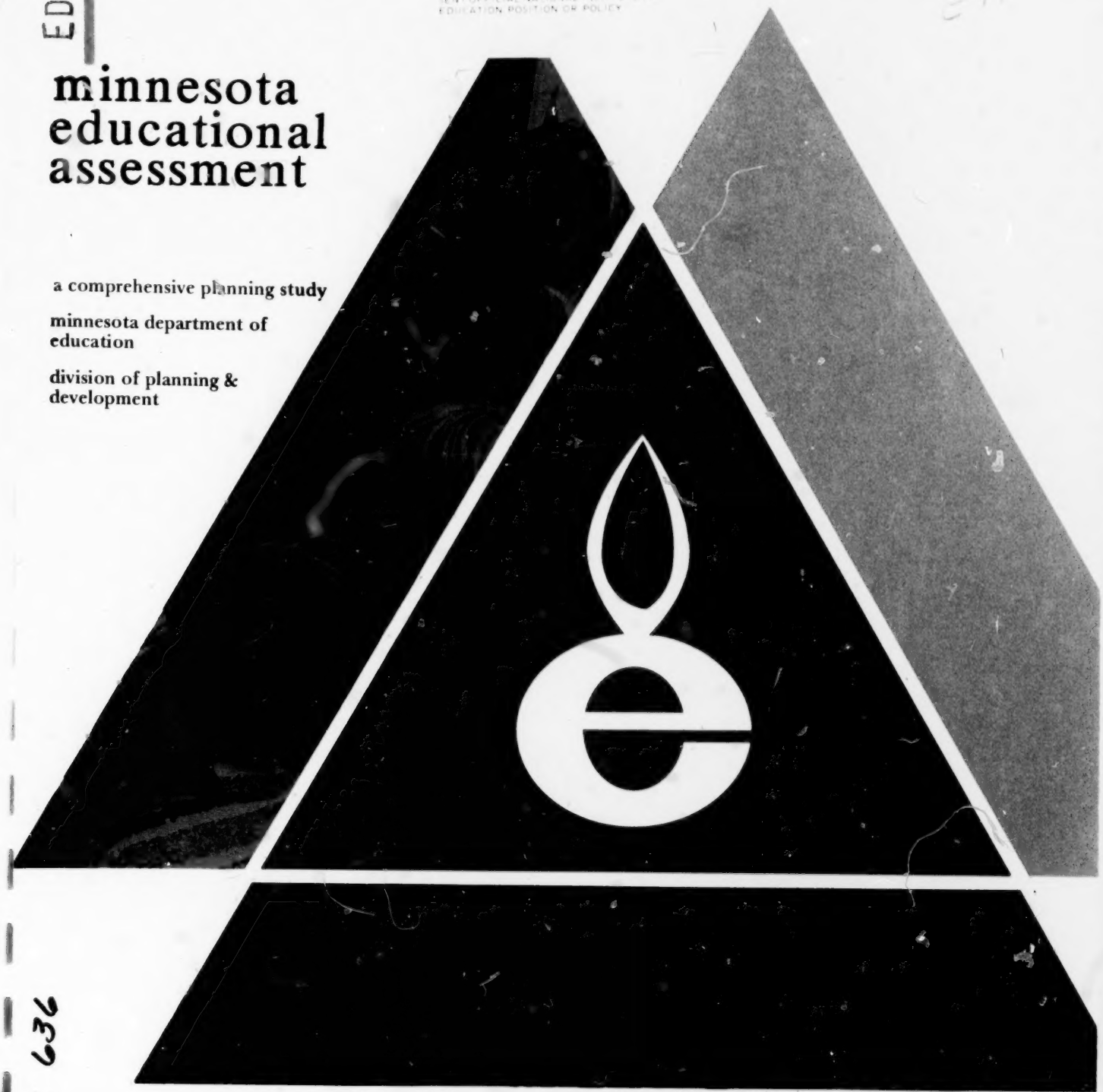
2415
1
EA

minnesota educational assessment

a comprehensive planning study

minnesota department of education

division of planning & development



EA005 636

ED 084657

MINNESOTA EDUCATIONAL ASSESSMENT: A COMPREHENSIVE PLANNING STUDY

Prepared By:
CENTER FOR EDUCATIONAL RESEARCH AND EVALUATION
Research Triangle Institute
Research Triangle Park, North Carolina 27709

Study Staff

DR. JOHN N. PYECHA	Senior Educational Systems Analyst Research Triangle Institute
DR. DAVID L. BAYLESS	Senior Educational Sampling Statistician Research Triangle Institute
DR. GEORGE H. DUNTEMAN	Senior Educational Statistician Research Triangle Institute
DR. JOHN P. BAILEY, JR.	Senior Educational Research Scientist Research Triangle Institute
DR. JOHN W. ADAMS	Director, State Educational Assessment Minnesota Department of Education

Research Triangle Institute

Research Triangle Park, North Carolina

January 1973

Prepared pursuant to regulations under Title III of
P.L. 89-10, as amended, of the Elementary and
Secondary Education Act of 1965

Minnesota Department of Education
St. Paul, Minnesota
Form X-A-9

ACKNOWLEDGEMENTS

The authors wish to express their special appreciation for the cooperation and contributions of the following personnel.

Commissioner Howard B. Casmey, Minnesota State Department of Education

Mr. Gregory J. Waddick, Assistant Commissioner, Division of Planning and Development

Dr. A. L. Finkner, Vice President, Research Triangle Institute

Dr. Frank B. Womer, Professor of Education, University of Michigan

Dr. Carmen Finley, Director of Assessment, Social and Educational Research Program, American Institutes for Research

Dr. Randall Johnson, Evaluator and Clinical Coordinator, Marshall-University High School and University of Minnesota

Dr. Mark Aulls, Assistant Professor, University of Minnesota

Members of the Minnesota Department of Education Assessment Advisory Committee:

Mr. David Dye

Dr. Carl Knutson

Dr. Gerald Kleve

Dr. G. Dean Miller

Mr. Wayne Erickson

Dr. Gayle H. Anderson

Mr. Len Nachman

Mr. Clyde Bezanson

Dr. Thomas Murray

Mr. Mel Johnson

Members of the Assessment Advisory Council and the Technical Advisory Committee to the Department of Education are listed in Appendix A.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	ix
LIST OF FIGURES	ix
Chapter 1: Executive Summary	1
I. INTRODUCTION	1
A. Purpose	1
B. Procedures	1
C. Organization of Report	2
II. SUMMARY	2
A. General Characteristics of the Proposed Assessment Program	2
B. Managing and Staffing (Chapter 3)	6
C. Instrumentation Development (Chapter 4)	8
D. Sample Design (Chapter 5)	9
E. Data Collection and Processing (Chapter 6)	10
F. Data Analysis (Chapter 7)	11
G. Reporting and Dissemination (Chapter 8)	13
Chapter 2: General Characteristics and Key Features of The Proposed State Assessment Program	15
I. INTRODUCTION	15
II. INTERPRETIVE LIMITATIONS OF STATE ASSESSMENT RESULTS	16
III. THE NAEP PROGRAM	18
A. What is National Assessment?	18
B. Why Follow the NAEP Model?	20
IV. CURRICULUM COVERAGE OF THE PROPOSED MINNESOTA PLAN	22
A. General	22
B. Grade and Age Level Coverage	22
C. Subject Matter Areas Covered	24
D. Testing Cycle	26
V. REPORTING VARIABLES	26
A. General	26
B. Candidate Variables	28
VI. THE PIGGYBACK OPTION	30
VII. DESIRED OUTCOMES	31
VIII. SUMMARY	33
Chapter 3: Management and Staffing	35
I. INTRODUCTION	35
II. MANAGEMENT AND STAFFING PLAN	35
A. Staffing Plan	35
B. Management	43

TABLE OF CONTENTS (continued)

		Page
III.	IMPLEMENTATION OF THE MANAGEMENT AND STAFFING PLAN	43
	A. The Phase 1 Assessment	43
	B. The Phase 2 Assessment	44
IV.	SUMMARY	45
Chapter 4:	Instrumentation Development	47
I.	INTRODUCTION	47
II.	GENERAL TIME FRAME AND STAFFING CONSIDERATIONS	47
III.	QUESTIONNAIRE SPECIFICATIONS AND DEVELOPMENT	50
	A. Nature and Purposes of Questionnaires	50
	B. Questionnaire Development	51
IV.	EXERCISE DEVELOPMENT	52
	A. General	52
	B. Definition of Behavioral Objectives (Task 1)	53
	C. Review of NAEP Materials (Task 2)	53
	D. Identification of Gaps (Task 3)	55
	E. Development of Exercises (Task 4)	55
	F. Packaging the Exercises (Task 5)	56
V.	SUMMARY	57
Chapter 5:	Sample Design	59
I.	INTRODUCTION	59
II.	GENERAL SAMPLE DESIGN CONSIDERATIONS, PRINCIPLES AND REQUIREMENTS	59
	A. Principle for Stratifying Schools	59
	B. The Distinction Between A Stratification and A Reporting Variable	60
	C. The Stages of Sampling and Principles for Sample Allocation	60
	D. The Concept of a "Design Effect"	61
	E. Requirements of the Sample Design	62
III.	Phase 1 SAMPLE DESIGN	63
	A. Introduction	63
	B. Stratification of the School Population	63
	C. Pupil and School Sample Allocations and Their Corresponding Statistical Precisions for the Reporting Regions	63
	D. Stratification Sample and Stratification of Pupils Within Selected Schools	69
	E. Summary of the Phase 1 Sample Design	69
Chapter 6:	Data Collection and Processing	71
I.	INTRODUCTION	71
II.	DATA COLLECTION	72
	A. General	72
	B. School Involvement	73
	C. Package Administration	74

TABLE OF CONTENTS (continued)

		Page
III.	DATA PPOCESSING	75
	A. General	75
	B. Data Editing	75
	C. Error Resolution	77
IV.	SUMMARY	78
Chapter 7:	Data Analysis Procedures	79
I.	INTRODUCTION	79
II.	PROCEDURES	79
	A. General	79
	B. Descriptive Analysis of Pupil and School Background Items	80
	C. Comparisons with National Assessment	81
	D. Within Minnesota Comparisons	81
III.	SUMMARY	84
Chapter 8:	Reporting and Dissemination	91
I.	INTRODUCTION	91
II.	PRE-ASSESSMENT WORKSHOPS FOR PHASE 1 (SPRING 1973)	93
III.	PHASE 1 (SPRING 1973) DISSEMINATION	93
IV.	PHASE 2 DISSEMINATION	95
V.	SUMMARY	96
REFERENCES		97
APPENDIX A		A-1

LIST OF TABLES

Table		Page
1	Curriculum Coverage of Proposed Minnesota Ten-Year Assessment Program	4
2	NAEP Subject Matter Assessment Plan for All Age Groups	19
3	Curriculum Coverage of the Proposed Minnesota Ten-Year Assessment Plan	23
4	Estimated Percentage of 9, 13, and 17-Year-Old Student Enrolled in Various Grades Across the United States	25
5	General Work Tasks for the Recommended Minnesota Assessment Program (3 Age/Grade Levels, 2 Subjects, 10 Reporting Regions)	36
6	Structure of the 17-Year-Olds Population for the Ten Reporting Regions	64
7	Alternative Sample Allocations and Their Statistical Precision for the State of Minnesota and the Ten Reporting Regions	65
8	Reporting Region School Sample Sizes and Number of School Strata for the Disproportional Allocation and the 17-Year-Old Population	68
9	Correct Response Rate (P-Values) by Reading Exercises for 9-Year-Olds in Minnesota, The Nation, and Central Region	85
10	Minnesota Assessment Results Versus NAEP National and Central Region by Reading Skill Areas and Age Levels	86
11	Correct Response Rate (P-Values) by Reading Exercises and School Size for Fourth Graders in Minnesota	87
12	Reading Results (Percent Correct Responses) in Word Attack Skills for Minnesota 17-Year-Olds by Race and Type of Community	88
13	Reading Results (Percent Correct Responses) for Minnesota 17-Year-Olds by Theme, Race, and Type of Community	89

LIST OF FIGURES

Figure		Page
1	Eleven Planning or Development Regions of the State of Minnesota	29
2	Management/Staffing Plan for Minnesota Educational Assessment Program	39
3	Milestones for Exercise Development in Region to Field Administration	49
4	The Minnesota Educational Assessment Sample Design	70

Chapter 1

Executive Summary

I. INTRODUCTION

A. Purpose

The Research Triangle Institute (RTI), under contract to the Minnesota Department of Education, assisted the Department in the design of a program for conducting periodic and comprehensive assessments of statewide educational progress. This is a report of that design study; the first chapter is both a summary and an introduction to the complete report.

The Minnesota Department of Education, prior to initiating this planning contract with RTI, planned and conducted a pilot educational assessment of a sample of students in grades 3 and 6. This pilot program involved an assessment of Reading, Mathematics, and selected areas in the affective domain during the 1971-72 school year. As a result of this pilot study, the Department decided to further explore alternative assessment approaches to fulfilling the following objectives that were adopted for the statewide assessment project:

1) To determine the level of performance of students in this state in the cognitive, affective, and psychomotor domains.
2) To identify the variables which are related to student performance.
3) To report the results of this investigation to educational decisionmakers in the Executive and Legislative branches of state government, the State Board of Education, the Department of Education, local school administrators, local school boards, and interested citizens of the state, thus providing a guide for the allocation of school resources.
4) To longitudinally report the extent to which progress is being made in Minnesota schools toward improving student performance within the State of Minnesota.

The planning approach for, and the results of, the pilot Minnesota assessment project are presented in references 1 and 2, respectively.

B. Procedures

This planning study was conducted in close liaison with the Division of Planning and Development, Minnesota State Department of Education. An iterative planning approach was used whereby alternative assessment plans were developed, costed, evaluated, and presented to the Department for consideration and feedback. This iterative process was repeated until an assessment plan was formulated to best meet the state's educational planning needs within the limitations of projected resources for supporting the program.

Each alternative plan was generally presented and evaluated in terms of six components or tasks that were identified as being essential to the design of an effective statewide assessment program:

1) Management and Staffing.
2) Instrumentation Selection and/or Development.
3) Sampling Design.
4) Data Collection and Processing.
5) Data Analysis.
6) Reporting and Dissemination.

Two excellent advisory groups established by the Minnesota Department of Education played important roles as feedback agents in the study procedure. These groups, an Assessment Advisory Council and a Technical Advisory Committee, represent an excellent cross-section of Minnesota's professional educators and lay persons who have an interest in the quality of the state's educational program. Both groups were extensively involved in this study; the Technical Advisory Committee provided feedback on technical considerations, the Assessment Advisory Council reacted to alternatives at a more general level. Members of both of these groups are listed in Appendix A.

C. Organization of Report

This report describes the assessment plan that is proposed for implementation in the State of Minnesota as a result of this planning study. Chapter 2 describes the general characteristics and key features of the plan. Chapters 3 through 8 contain methodologically oriented discussions of the six essential components of the plan as follows:

Chapter 3, Management and Staffing.
Chapter 4, Instrumentation Development.
Chapter 5, Sample Design.
Chapter 6, Data Collection and Processing.
Chapter 7, Data Analysis.
Chapter 8, Reporting and Dissemination.

An overview of these chapters is provided in the summary section below.

II. SUMMARY

A. General Characteristics of the Proposed Assessment Program

The proposed educational assessment plan represents a vehicle whereby Minnesota can conduct a comprehensive assessment of the state's educational progress.

It can provide a means of periodically monitoring achievement in the cognitive, effective, and psychomotor domains.

The assessment program cannot, however, demonstrate or show causes for differences in achievement between various groups of students. Assessment results serve only to describe the achievement status (by Reading or any other output variable) of selected groups of students at the time the measures were taken. As such they serve to spotlight achievement discrepancies between selected groups of students. Additional state resources (generally as a research function) can then be focused on "why" such discrepancies do in fact exist and "what" can be done about them. Nothing but experimentation, if that, can serve to demonstrate the causes of these "between group" discrepancies in achievement.

The proposed Minnesota Educational Assessment program is modeled after the National Assessment of Education Progress (NAEP), an ongoing educational project designed to give educators and the lay public a better look at those knowledges and skills that American youth have acquired. NAEP provides for a systematic, continuous, census-like survey of knowledges, skills, understandings, and attitudes as exhibited by students and young adults in four age levels and across ten different subject areas. By following the NAEP model, Minnesota can: (1) reduce the costs of developing and scoring assessment exercises; (2) compare the assessment results for Minnesota students to NAEP results for students in the Nation as a whole, as well as in the Central Region* of the Nation (all comparisons will be made at the same student age levels); and (3) take advantage of past and future exercise administration, data collection, sampling, and data analysis methodologies that have been (and will continue to be) developed by NAEP.

Table 1 presents an overview of the curriculum coverage provided by the proposed Minnesota plan. (Curriculum coverage is defined to include the grade and age levels assessed, the subject matter areas included in the assessment program, and the reassessment cycle for measuring educational progress.) As shown in Table 1, the plan provides for the collection of assessment data from a sample of students in grades 4, 8, and 11, and from a sample of students who are in school and are 9, 13, and 17 years of age. However, grade and age samples will overlap considerably. For example, over 70 percent of the 9 year olds will be 4th grade students. The age level results will be used for making Minnesota versus NAEP comparisons; the grade level results will be used to enhance state level decisionmaking. Assessment data will be collected from 8th graders and from 13-year-olds in October and November, from 4th graders and 9-year-olds in January, and from 11th graders and 17-year-olds in March and April.

Subject area coverage would be initially limited to the ten NAEP subject matter areas as shown in Table 1; i.e., Art, Career and Occupational Development, Citizenship,

---

* The Central Region includes the following states: North Dakota, South Dakota, Iowa, Kansas, Minnesota, Missouri, Nebraska, Illinois, Michigan, Ohio, and Wisconsin.

Table 1

CURRICULUM COVERAGE OF PROPOSED MINNESOTA TEN-YEAR ASSESSMENT PROGRAM

Year	Curriculum Coverage By Grade/Age Levels		
	4th Graders and 9-Year-Olds[1/]	8th Graders and 13-Year-Olds[1/]	11th Graders and In-School 17-Year-Olds[1/]
Cycle 1:			
1972-73			Reading
1973-74	Reading & Literature	Reading & Literature	Literature
1974-75	Math & Social Studies	Math & Social Studies	Math & Social Studies
1975-76	Music & Science	Music & Science	Music & Science
1976-77	Writing & Career and Occupational Development	Writing & Career and Occupational Development	Writing & Career and Occupational Development
1977-78	Citizenship & Art	Citizenship & Art	Citizenship & Art
Cycle 2:			
1978-79	Reading & Literature	Reading & Literature	Reading & Literature
1979-80	Math & Social Studies	Math & Social Studies	Math & Social Studies
1980-81	Music & Science	Music & Science	Music & Science
1981-82	Writing & Career and Occupational Development	Writing & Career and Occupational Development	Writing & Career and Occupational Development
1982-83	Citizenship & Art	Citizenship & Art	Citizenship & Art

1/ Eighth graders and 13-year-olds are tested in October-November; 4th graders and 9-year-olds are tested in January; and 11th graders and in-school 17-year-olds are tested in March-April. Note also that 9-year-olds will generally be distributed across grades 2-5; 13-year-olds across grades 5-9; and 17-year-olds across grades 9-12.

Literature, Mathematics, Music, Reading, Science, Social Studies, and Writing. Those exercises which were developed in these subject areas by NAEP and which are relevant to Minnesota educational objectives would be used by Minnesota. Supplementary exercises would be developed by Minnesota to assess the state objectives not covered by these NAEP exericses. Although these ten subject areas do provide for comprehensive curriculum coverage, it is recommended that Minnesota expand this coverage, as resources permit, to include Health, Physical Education, and measures in the affective domains. The curriculum coverage plan of the proposed assessment program has the flexibility to be expanded, reduced, and otherwise modified as future events dictate.

Except for the initial assessment phase scheduled for the spring of 1973 (Reading for 11th graders and in-school 17-year-olds), two subject areas are scheduled to be assessed during each school year. The spring of 1973 assessment in Reading has been designated as Phase 1 of the program, the subsequent ten years of the assessment program as Phase 2. This modified schedule for the Phase 1 spring assessment stems from the desire of the Department of Education to maintain the assessment thrust of the pilot assessment project. Although the Department can develop the necessary supplementary Reading performance items for the Phase 1 spring assessment, they cannot develop the Literature items in time. However, this modification in the Phase 1 assessment has only a minor effect on the reassessment testing cycle. The program is thus based on a five-year "reassessment" cycle. That is, each subject area is reassessed in five years (except for the first six year cycle for Reading of 11th graders and in-school 17-year-olds). In addition to allowing a reasonable time for measuring educational progress, this cycle provides time for the preparation of exercise items for the reassessment.

The utility of the assessment results depends upon how groups of students are defined for reporting purposes. State results for each of the three age groups would be used to report comparisons between Minnesota students and students in the NAEP National and Central Region samples. Within state results would be reported by additional groupings of students within each of the three grade levels (4, 8, and 11). A final delineation of these reporting groups is pending an exact determination of funds available for the Phase 1 assessment. However, a list of candidate reporting variables was developed and ranked by order of importance. This list provides for grouping student results by: (1) such student related variables as sex, SES, race, longevity in the Minnesota Educational System, and the type and degree of participation in extra-curricular activities; (2) such school related variables as public and/or non-public status, type and size of community in which school is located, size of enrollment, teacher/student ratio, per pupil expenditure levels, regional location,

and selected characteristics of staff and student body; and (3) process related variables pertinent to instructional approach and/or program participation.

The proposed Minnesota assessment plan has two unique features that could be introduced on a pilot basis during the Phase 1 assessment. The first is an option whereby local school districts could linkup with, or "piggyback" onto, the state assessment to obtain results for their districts. The Department of Education, though it has endorsed this option, has decided to field test it with one or two districts before finalizing procedures and policies for its full implementation. Plans at this time anticipate that local school districts would be responsible for any additional costs associated with exercising this option.

The second unique feature is the development of "desired outcomes" measures that would reflect desired performance levels for Minnesota students in the assessed subject matter areas. This judgmental information could assist educational decision-makers in identifying areas of concern and commendation within the state. There are, however, several technical problems involved in developing these measures. As a result, implementation of this feature would be on an experimental basis, as limited by the availability of resources.

B. Managing and Staffing (Chapter 3)

The ten-year educational assessment program proposed for Minnesota involves the annual collection, analysis, and dissemination of a substantial amount of educational data. That is, with the exception of the first and last years of the assessment cycle, Department of Education personnel will be simultaneously involved in collecting data for the current year's assessment, analyzing and disseminating the results for the previous year's assessment, and preparing for next year's assessment (e.g., developing exercises and finalizing the data collection strategy). The recommended program, by following the NAEP model of collecting data on 13-year-olds in the fall (October-December), on 9-year-olds in the winter (January-February), and on in-school 17-year-olds in the spring (March-April), does serve to reduce the data collection work load by distributing it over the school year.

The current staff of the Minnesota Assessment Project consists of a Director and typing assistance, plus two excellent advisory groups (the Assessment Advisory Council and the Technical Advisory Committee). However, an assessment program of this magnitude entails a wide range of work tasks, many of which require highly specialized technical and professional skills. That is, Minnesota would need a full-time staff of nine professionals, two junior professional/research assistants, one secretary, two clerk/typists, and four typists, plus eight part-time exercise administrators and some specialized contracting and/or consulting assistance to implement and sustain the proposed assessment program. (It is not always cost-effective to hire specialized expertise in-house, especially when this expertise

is required in brief, infrequent intervals throughout each assessment year. Hence, the additional contracting and/or consulting services would be used in exercise development, sampling, data analyses, and exercise scoring.)

The full time Minnesota assessment staff, contracting/consulting services, and advisory groups would be organized as follows:

1) Director's Office: A Director and a Secretary.
2) Advisory Groups: An Assessment Advisory Council and a Technical Advisory Committee.
3) Instrumentation Section: A Head, a Research Assistant, two Typists. and a subcontracted team of six Exercise Writers.
4) Survey Operation Section: A Head, a Clerk/Typist, two District Supervisors, and eight part-time Exercise Administrators.
5) Scoring, Data Reduction, Sampling, Data Analyses, and Report Preparation Section: A Head, two Typists, an Educational Research Analyst, an Educational Sampling Statistician, and a Research Assistant. This Section would also require contracted/consultant assistance for scoring open-ended exercises, for designing the sample and data analysis plans, and for computerized scoring and data reduction services.
6) Dissemination Section: A Head and a Clerk/Typist.

These personnel would be initially assigned 100 percent of their time to the assessment program. However, by coordinating assignments to coincide with slack periods in the annual program, staff members could assist other Divisions within the Department in designing and conducting research studies and/or surveys.

It would be neither possible nor desirable for the Minnesota Department of Education to recruit and/or develop the above staff in time for the Phase 1 assessment. Hence, a strategy for implementing the Management and Staffing Plan for the Phase 1 and Phase 2 assessments was developed. This strategy is based on a conservative, graduated approach whereby key personnel would be hired, given on-the-job training, and gradually phased into the program over a period of one or two years. This approach would initially require contracting out most of the work tasks; however, responsibility for these tasks would be gradually assumed by the Department of Education as key personnel are moved into the program. This recommended strategy provides flexibility for handling uncontrollable events that might arise, while maintaining continuity in the program. Furthermore, this approach would not place a newly organized and relatively inexperienced staff in the formidable position of having to implement an assessment program of this magnitude.

C. Instrumentation Development (Chapter 4)

Two types of instrumentation are required for the proposed educational assessment program. The first is the exercises for measuring performance in subject areas; the second is the questionnaires for obtaining student and school background information. If the desired outcomes measure is also included as a component of the assessment program, the instrumentation requirments would be expanded to include additional questionnaires for gathering desired outcomes information.

NAEP exercises are released in December or January of the school year following their use in the National Assessment program. With few exceptions, the Minnesota curriculum coverage plan is designed to incorporate the released NAEP exercises into the Minnesota assessment program, beginning in the fall of the school year following their release.

Given these constraints, the exercise development year for each pair of subject matter areas extends roughly from November of one year to December of the next year. At least two months is required after the exercises are developed for packaging, printing, and distributing the assessment materials in time for their scheduled fall, winter, and spring administrations.

The plan for identifying and developing the exercises required to supplement the released NAEP items involves five major tasks:

1) Definition of broad state goals and their translation into measurable, operational and behavioral objectives.
2) Review of NAEP objectives and released exercises in order to select those exercises which are relevant to Minnesota objectives.
3) Identification of the gaps for those Minnesota objectives which are not adequately measured by the selected NAEP exercises.
4) Development of required exercises.
5) Packaging of the exercises into booklets.

The magnitude of the annual exercise development task would vary depending upon the degree to which Minnesota objectives are measured by the released NAEP exercises (as indicated by the results of Tasks 2 and 3 above). NAEP, on the basis of five years of experience in developing their exercises, has formulated a revised approach to this critical task (Task 4 above). A modification of this revised NAEP plan is recommended for the proposed Minnesota educational assessment. NAEP contracts most of the exercise development. It would be less expensive for Minnesota to organize, train, and supervise working groups of school personnel to assist with exercise development. These personnel would be supplemented by local consultants from the University of Minnesota or state colleges and by consultants knowledgeable about the development of NAEP exercises. The composition of the exercise development team,

which would generally begin its 12-month-task in December, would vary annually as different subject matter areas are rotated into the program. Hence, the recommended plan would generally follow NAEP procedures, but would require less outside contractual assistance--especially after the program has been in operation for a few years.

An important consideration in completing all five of the above tasks is the involvement of subject specialists, educators, lay people, and, perhaps, students in reviewing the objectives and exercises at key stages in their development. Reviews of exercises should take into account their potential offensiveness, ease of scoring, ease of administration, content validity, and content appropriateness. Using a form of matrix sampling, these exercises could be distributed over several exercise packages to minimize the exercise administration time per student; i.e., each student in the sample could take only a portion of the total number of exercises involved in the assessment.

Major steps in the questionnaire development plan include:

1) A delineation of the variables and data elements to be included in the questionnaires.
2) A review of available questionnaires for items that can be incorporated into the required questionnaires.
3) The development of draft questionnaires.
4) A field test(s) to tryout the questionnaires.
5) A revision and finalization of the draft questionnaires.

As special considerations, all questionnaires should be designed to be brief, amenable to transformation to a computerized file by optical scanners, and compatible with the data collection and analysis plans.

D. Sample Design (Chapter 5)

Sampling, which provides educational decisionmakers and those interested in education with results of sufficient precision at a reasonable cost, was selected over the alternative of collecting assessment data from every Minnesota student in the grade/age levels to be assessed.

In general, the sample design recommended for Minnesota meets the following requirements:

1) The sample is a probability sample; i.e., each student in a given age class or grade in a public or non-public school in the State of Minnesota has a known positive chance of inclusion in the sample.
2) Each of the ten geographical reporting regions of the state (two of the eleven regions are combined to form a single region) are represented in the sample so that results can be reported for each of them with nearly equal statistical precision.

3) Each reporting variable, or any combination of two reporting variables, can be reported for up to eight reporting groups. This requirement allows for the analysis of interaction effects of certain pairs of reporting variables; i.e., differences in the effects of one reporting variable for different levels of the remaining variable will be examined.

4) If, for a given grade and/or age level, the amount of time it takes to administer the entire set of assessment exercises to a given student is longer than desirable, a matrix sampling approach will be developed to shorten the length of time each pupil will be tested.

5) The school and pupil school sample sizes will be such that estimates of the sampling variability of the reported results can be estimated from the sample data.

This sample design involves the use of the state's ten reporting regions as a stratification variable to give the sample "representative credibility" by geographically spreading it across the state. Selected school level variables which relate to educational performance are also used as stratification variables within each of the ten reporting regions. Stratifying by these achievement related school variables (per pupil expenditure and type and size of community in which the school is located) provides for more homogeneous groupings of schools, thus enhancing the statistical precision of assessment results.

The sample design involves two sampling stages. The first stage would consist of selecting a random sample of schools within strata; the second stage would be the selection of a random sample of students within the schools selected at stage one.

Using this sample design, the statewide probability sample for the Phase 1 assessment of in-school 17-year-olds would include approximately 5,100 students from 270 schools. The Phase 1 assessment of Reading would involve no more than two hours of total exercise administration time. Consequently, it will not be necessary to use a matrix sampling approach in this phase of the assessment.

E. Data Collection and Processing (Chapter 6)

The quality of information gathered is greatly affected by how it is collected and who collects it. Since good decisions are rarely made on the basis of poor information, the task of collecting and processing data (includes editing and scoring) constitutes an integral aspect of the assessment plan.

A field survey approach analogous to that used in the ongoing NAEP program is recommended as the most cost-effective way for Minnesota to collect their assessment data. Specially trained survey teams, using a "one day in--one day out" approach, will administer exercises and collect background information on the students and schools

in the statewide sample. Data collection at each school will require no more than one-half day and will be conducted with minimal disruption to school programs and slight impositions on students, teachers, principals, and other school officials.

These survey teams of District Supervisors and Exercise Administrators will be supervised and coordinated by the Head of the Survey Operations Section. Exercise Administrators will use paced tapes to better ensure the standardization of all data collection and, except when assessing Reading, to help prevent the exercises from measuring reading ability as well as the subject area being assessed.

Assistance and cooperation of each school selected in the sample will generally be required as follows:

1) To provide a roster of eligible students in the proper grade and age levels (this information would be used to draw the student sample).
2) To designate one school official, preferably a counselor, to help arrange the testing schedule and to aid the District Supervisor and Exercise Administrators in having students at the administration site on schedule.
3) To provide adequate space for administering the exercises (exercise packages will be administered to an average of approximately 20 students from each school in the sample).
4) To cooperate in completing the school questionnaires and in providing some of the background information for elementary school students selected in the sample.

Assessment data will be collected on machine readable/scorable forms in order to minimize errors inherent in reproducing or transferring data to magnetic tape files for computerized retrieval. However, editing and error resolution activities that require various levels of judgment and involve more than one individual will also be conducted at several phases in the data handling process.

Because it is important to safeguard the rights of participating students and school principals, certain measures are included in the data collection and processing plan to assure confidentiality with respect to all information collected on individual students and schools.

As such, these data collection and processing procedures assure a high degree of cooperation and a great degress of quality control with respect to both sample selection and data collection.

F. Data Analysis (Chapter 7)

The data analysis plan for the proposed assessment program focuses on various strategies of statistical analysis to extract the most important and relevant descriptive measures from the data, as well as to detect important differences in achievement measures between various subgroups of students.

This data analysis plan consists of three general sets of analyses. The first is a descriptive analysis of the responses to the student and school questionnaires that would serve to describe the input characteristics of the students and schools in Minnesota. That is, estimated proportions of Minnesota students in each grade/age level who are in each of the discrete categories of the selected reporting variables would be computed; e.g., numbers of students attending schools in Large City, Large City Fringe/Medium City, and Small Town/Rural types of communities. Similar estimates would also be computed for the proportion of Minnesota schools in all reporting groups for the school related variables; e.g., numbers of schools with various average student/teacher ratios. In addition, estimated proportions would also be compiled for various combinations of reporting variables.

The second set of analyses involves comparisons of Minnesota Assessment results to National Assessment results. That is, the exercise p-values for various groups of Minnesota students would be compared to those of students in comparable age groups (9, 13, and in-school 17-year-olds) from throughout the Nation and Central Region who participated in the National Assessment Program. (A p-value is a statistic estimating the proportion of students who respond correctly to an assessment exercise.) Such comparisons can only be made on these released NAEP exercises that are adopted for use in Minnesota.

The third and final set of analyses involves the computation and comparisons of assessment results for various groups of students within Minnesota as defined by the reporting variables. Two levels of analysis would be conducted in this phase of the analysis plan. The first level is primarily descriptive in orientation and would involve the calculation of exercise p-values, their respective standard errors, and certain comparisons involving pairs of p-values. At the second level of analysis, relationships between exercise responses and selected student and school background variables would be investigated by multiple regression, balanced fits, and certain multivariate techniques.

These analytical techniques involve the simultaneous examination of a number of independent variables and are useful for examining the effects of one independent variable while controlling or adjusting for the effects of other independent variables. These second level analyses are directed to investigating the relationships between achievement measures (exercise responses) and selected student, family, and school characteristics. One might, for example, pose the following question under this level of analysis: What would be the differences in p-values between students in schools with large enrollment differences if the distributions of students by parental education, race, SES, type of community in which the school is located, etc., had been

the same for all schools? A great danger in this approach is attributing casualty to those variables that turn out to be significant in the statistical model. Nothing short of experimentation, if that, can demonstrate what the actual effects of these variables are. Statistical models based on survey data (non-experimental data) can only provide valuable hints and insights into the probable effects of these variables.

G. Reporting and Dissemination (Chapter 8)

The basic purpose of a dissemination plan for statewide assessment results is to insure that accurate information is made available to all interested people in the state, at a level of sophistication (detail) commensurate with their background, needs, and purposes. This information is needed so that people at all levels can evaluate properly the need for change, so that they have enough information of the right kind and type to make intelligent and data-based educational decisions. Thus a legislator or high-level policymaker would have a far different background, need, and purpose for assessment results than would a housewife with children in school. Too little information in the former case would be deleterious, while too much in the latter case would provide unneeded, unwanted, and perhaps misunderstood data.

There are at least four levels which should be considered in understanding the various needs of the consumers of assessment program results. The policymaking level consists of legislators and members of the executive branch of the state government who are charged with the responsibility for establishing broad policies which, when carried out, will best meet the educational needs of the state. A decisionmaking level is comprised of personnal throughout the state, and in the Department of Education, who are charged with carrying out policy through making the basic operational decisions and allocations of resources. An operational level or the "on the street" level is one in which educators work actively with children and parents, face the very real problems in modern education, and carry out all the policies and decisions which have been made. And, finally, a public level brings us back full circle to the policymaking level. The public level consists of local school board members, various special interest groups, both concerned and somewhat-less-than-concerned parents of children in school, and a very large group of adults with no children in school. This group, through the democratic processes, can and should have a strong effect on the first level.

The dissemination plan must be responsive to the informational needs of these various levels of consumers of educational assessment results and of educational decisionmakers who will use assessment information. All levels must have access to all results; but their specific information needs vary. This means, then, that report formats, audio-visual aids, presentation modes, and any other form of promulgation and publication decided upon and used should be tailored to some degree to decisionmaking groups, special interest groups, and the general public.

A series of pre-assessment workshops is proposed in spring 1973, before actual Phase 1 assessment administration, to acquaint the decisionmaking and operational levels with the project. In addition, wide public exposure to the assessment project is suggested through television, service clubs, and special interest group presentations.

The primary vehicle for assessment dissemination is seen as a series of Fall Workshops, each of which serves to disseminate the results of the previous school year's assessment and to introduce future assessments. Actual dissemination of Phase 1 results will occur by means of the basic Technical Report, a widely distributed Highlights Report that would be written in popular language and format, and Fall Workshops for decisionmaking and operational persons in the subject area(s) assessed. These workshops will also serve to introduce the early Phase 2 assessments.

Yearly Phase 2 dissemination will be in the same manner: Technical Report, Highlights Report, and dissemination Fall Workshops on subject-matter areas. Each year there will also be special presentations prepared for the policymaking level.

Chapter 2

General Characteristics and Key Features of The Proposed State Assessment Program

I. INTRODUCTION

This chapter provides a capsule description of the general characterists and key features of the proposed Statewide Assessment Program that are not covered in the subsequent, more methodologically oriented chapters of this report.

During the course of this study, the authors had the opportunity to interact with, and solicit the opinions and counsel of, an excellent cross section of the Minnesota citizenry (legislators, professional educators, representatives of special interest groups, leaders among local schools, the lay public, etc.) who have an interest in and a concern for education throughout the State. In pursuing the topic of state assessment with these people, in group or individual discussions, it became apparent that a general misunderstanding exists as to the nature of the information provided by a statewide assessment program. This chapter begins, therefore, with a brief discussion relative to the types of information provided by a state assessment and some cautions as to how this information can be interpreted.

The ongoing National Assessment of Educational Progress (NAEP) project is the next topic considered in this chapter. Since the proposed Minnesota program is based on the NAEP model, this section provides both an overview of NAEP and a brief discussion of the major factors influencing Minnesota's decision to follow the NAEP model.

The curriculum coverage provided over the next ten years by the assessment program is then discussed in this chapter in terms of which age and grade levels are assessed and when, the subject matter areas included in the program, and the reassessment cycle whereby the educational progress of the State can be evaluated.

The utility of the state assessment for educational decisionmaking depends to a large extent upon how various groups of students within the state are delineated for the purpose of reporting the assessment results. A general strategy for delineating these reporting variables, along with a candidate list of student, school, and process reporting variables, are also presented in this chapter.

Additional features of the proposed assessment program that are covered in separate sections of this chapter include an option whereby local school districts can link up with or "piggyback" onto the state program to provide assessment results for their districts, and an experimental feature for exploring the development of performance criteria or desired outcomes against which state assessment results could

be compared as one means of identifying areas of concern and commendation within the state educational program.

II. INTERPRETIVE LIMITATIONS OF STATE ASSESSMENT RESULTS

A statewide assessment program is an important component of the state department planning function. Properly conceived, the statewide assessment program can provide a means of periodically monitoring achievement in the cognitive, affective, and psychomotor domains to determine whether or not Minnesota children know and can do those things that they should be able to do in order to live a full life.

In addition to providing status reports of performance with respect to desired outcomes at various stages in the student's scholastic career, the assessment program can serve to identify those groups of students, by certain general characteristics (e.g., sex, race, SES, and geographic location and resource levels of schools attended), who may or may not be realizing the educational objectives of the state. (This is an important factor if one adheres to the belief that the state should be accountable to all children, regardless of their race, geographic location, and family status.) Given such information, educators can formulate plans for the allocation of resources so that the attainment of desired outcomes can be enhanced among those groups of students having the greatest needs.

However, it must be noted that assessment results generally serve only to describe the achievement status (by Reading or any other output variable) of selected groups of students at the time the measures were obtained. As such, they serve to spotlight achievement discrepancies between selected groups of students. Additional state resources (generally as a research function) can then be focused on "why" such discrepancies do in fact exist and "what" can be done about them. The assessment results themselves DO NOT SHOW CAUSAL RELATIONSHIPS. That is, differences in achievement for different categories of a reporting variable such as size of school could also be due to differences in other school variables (e.g., student-teacher ratio) and/or student characteristics (e.g., aptitude) among the different categories of size of school. Consequently, one cannot be reasonably sure that differences in achievement are due to differences in size of school. In order to make a causal statement, one must have had the opportunity to allocate children randomly to the various size of school categories--a procedure that is not generally feasible nor desirable.

Most decisionmakers are concerned with the costs of education. Unfortunately, however, cost/effectiveness studies in education are extremely complex and, given present school accounting procedures, statewide assessment programs can provide only limited "cost/effective" information. Furthermore, reporting by fiscal categories

is especially conducive to misinterpretation. For example, one might find little variation in Reading achievement among students in schools having a wide range of per pupil expenditures. One might quickly conclude that increasing educational expenditures is not likely to result in improved Reading. However, the between school variation reflected in the overall per pupil expenditures might not be a good measure of the between school variation in Reading instruction expenditures. That is, observed "between school variations" may have been caused by great variations in expenditures for space, supplies, and equipment in such areas as Science, Biology, or Physical Education; whereas between school variations in Reading per pupil expenditures were very small.

A good general example of one of the problems involved in trying to make causal inferences from survey data is provided by the following portion of a "newsnote" that appeared in a recent issue of Phi Delta Kappan. [Ref. 3]

> Most schoolmen regard Roger Freeman as one of the more mendacious authors on school finance in the U.S. A senior fellow at the Hoover Institution of War, Revolution, and Peace at Stanford University, Freeman said in a Wall Street Journal article last March 31 that "the higher the expenditures per pupil--the smaller the class size--the lower are pupil achievements--and vice versa."
>
> Citing as his source the New York City School Fact Book, Freeman announced that in 1967068 there were 30 New York schools in which per-pupil expenditures averaged $1,330. Then there were 101 schools in which the average was $441. Thus the first group of schools spent about 2.5 times as much as the second. The teacher-pupil ratio was 1:12.3 in the high-expenditure schools, 1:25.9 in the low-expenditure schools--or more than twice as high.
>
> But the reading skills of the students in the low-expenditure, large-class schools averaged above grade level, Freeman said, while in the high-expenditure, small-class schools they were below grade level. "This is not just an accident," Freeman alleged. "A review of . . . reports from other cities shows that the high-expenditure, small-class city school typically is one with low educational achievements."
>
> Albert Shanker says in his weekly New York Times column, "Where We Stand," that Freeman's views would merit no more than casual attention if they were the views of one man; but the ominous fact is that these dangerous views are the hallmark of an odd coalition of the right and the left, the wealthy right embracing them as a justification for withholding adequate fiscal support for education, the new left proclaiming that "relevancy," "Community control," "ethnic studies," and "life-style" rather than more money are the keys to educational success.

"The shoddiness of such reasonings is quite apparent," Shanker says. "Pupils do not do poorly in reading and math because they are in small classes; rather, they are placed in small classes because they are doing poorly. If we were to extend Freeman's logic to [medicine], the more money an individual spends on doctors and hospitals, the poorer his health. The healthiest individuals spend little or no money on doctors and hospitals. Hence the way to fight disease is to abolish Medicare, health insurance, and welfare programs."

Shanker suggests that even though some cities are spending more than others, they are not spending enough. Or, he adds, it may be that, while more money will make some difference, the schools, no matter how effectively they function, cannot overcome all the nonschool factors which prevent students from achieving. . . .

III. THE NAEP PROGRAM

A. What is National Assessment?

This ongoing educational project is designed to give educators and the lay public a better look at those knowledges and skills that American youth have acquired. The NAEP plan provides for a systematic, continuous, census-like survey of knowledges, skills, understandings, and attitudes as exhibited by students and young adults in four age levels and across ten different subject areas. The ultimate goal of National Assessment in providing this information is to improve the educational process, to improve education at any and all levels where knowledge will be useful about what students know, what skills they have developed, or what their attitudes are. [Refs. 4 and 5]

The four age groups used in NAEP are 9, 13, and 17-year-olds who are in school, 17-year-olds who are not in school, and young adults in the 26-35 age range. All age groups are assessed annually. The ten subject areas assessed are Art, Career and Occupational Development, Citizenship, Literature, Mathematics, Music, Reading, Science, Social Studies, and Writing. The NAEP subject matter assessment plan is given in Table 2. Note that two subject areas are assessed annually (except for the 1969-70 assessment) and that the reassessment of subject areas in a five-year cycle allows for comparisons to show whether change has occurred.

The results for about fifty percent of the NAEP exercises given each year are reported--for each exercise and each age group--by the following categories (beginning with the second assessment year, results are reported for each student

Table 2

NAEP SUBJECT MATTER ASSESSMENT PLAN FOR ALL AGE GROUPS*

Cycle 1

March 1969 - February 1970	Science, Writing, Citizenship
October 1970 - August 1971	Reading, Literature
October 1971 - August 1972	Music, Social Studies
October 1972 - August 1973	Math, Science
October 1973 - August 1974	Writing, Career and Occupational Development
October 1974 - August 1975	Citizenship, Art.

Cycle 2

October 1975 - August 1976	Reading, Literature
October 1976 - August 1977	Music, Social Studies
October 1977 - August 1978	Math, Science
October 1978 - August 1979	Career and Occupational Development, Writing
October 1979 - August 1980	Citizenship, Art

* Source: Questions and Answers About the National Assessment of Education Progress. Denver, Colorado: The National Assessment of Educational Progress, Education Commission of the States, April 1972, p. 5.

area by theme, i.e., a set of exercises which share a common content but which may require diverse behavioral responses):

1) Gegoraphic region--Northeast, Southeast, Central and West.
2) Size and Type of Community--extreme rural areas, extreme inner cities, extreme affluent suburbs, inner city fringes, suburban fringes, medium cities, small cities.
3) Sex.
4) Color.
5) Parental education--both parents with eighth grade or less, at least one parent with some high school but not graduated, at least one parent graduated from high school, and at least one parent with some post-high school training. [Ref. 5]

These "reported" exercises are then released and can be used in state and/or local assessment programs.

Additional features of the NAEP model which are also incorporated into the Minnesota Assessment Program include the use of probability sampling and the use of trained exercise administrators and paced tapes to standardize the collection of assessment data. Detailed descriptions of these and other general features of the NAEP model are provided in References 3, 4, and 5.

B. Why Follow the NAEP Model?

After considering alternative general assessment models, three basic factors influenced the Department of Education's decision to follow the NAEP model, or some variation thereof.

1. Reduced Exercise Development and Scoring Costs

The Department made an early decision to forego the use of norm-referenced standardized tests in favor of the criterion-referenced approach. However, the development of criterion-referenced items can be extremely costly. As indicated above, NAEP releases approximately fifty percent of their exercises and these released exercises could be used in state and/or local assessment programs. Hence, Minnesota can cut down on exercise development and scoring costs by working with the released exercises that were developed at NAEP expense to cover a wide range of subject areas.

The subject matter objectives measured by the NAEP exercises have to be acceptable to three groups of people: [Ref. 6]

a) Subject Matter Specialists. Specialists in the subject area must consider the objectives authentic from the viewpoint of the discipline; scientists must agree the Science objectives are authentic, mathematicians must agree upon the authenticity of the Mathematics objectives, etc.

b) Educators. School people must recognize the objectives as being desirable for education and as being the types of objectives which schools are actively striving to achieve.

c) Citizens. Parents and others interested in education must agree that the objectives are important for youth and young adults to know, feel, or understand.

In the development of all objectives as per these stringent criteria, NAEP thus takes an important step that is not commonly undertaken by educators; that is, lay people are extensively involved in reviewing all objectives.

Given these subject matter objectives, NAEP exercise writers are directed:

a) To develop exercises in whatever form or mode deemed most appropriate to the assessment of a particular objective.

b) To develop exercises that are samples of some important knowledge, or skill or attitude.

c) To develop exercises that sample equally those attributes common to most of those assessed, to about half of those assessed, and to the ablest, most knowledgeable assessed of a given age. [Ref. 6]

Despite this thorough and comprehensive NAEP approach to exercise development, the Department of Education should exercise caution against assuming that all the NAEP exercises are "right for Minnesota"--NAEP objectives and their corresponding exercises should be examined for relevance to the Minnesota objectives. Following this review, irrelevant objectives and/or exercises should be dropped and additional objectives and/or exercises developed to fill in the gaps.

2. Minnesota Versus NAEP Assessment Comparisons

Linkage to the NAEP model provides two important external reference points for Minnesota educators; i.e., comparisons to National results and comparisons to Central Region* results. It must be noted that these comparisons are age level and not grade level comparisons--NAEP does not yet collect data and report results by grade levels. As discussed in the next section, Minnesota statewide data will be collected, and results will be reported, by three grade levels (4, 8, and 11), in addition to three NAEP age levels (9, 13, and 17-year-olds; Minnesota will not assess the 26-35 age group). Thus, Minnesota can use the age level comparisons as general reference points, while using grade level results as the basis for statewide educational planning.

---

* The Central Region includes the following states: North Dakota, South Dakota, Iowa, Kansas, Minnesota, Missouri, Nebraska, Illinois, Michigan, Ohio, and Wisconsin.

3. Carryover of NAEP Technology

Minnesota, by following the NAEP model, can take advantage of the test administration, data collection, sampling, and data analysis methodologies that have been developed, field tested, and perfected through the years by NAEP. Since NAEP is an ongoing project, further relevant NAEP innovations can likewise be adopted by the Minnesota program.

IV. CURRICULUM COVERAGE OF THE PROPOSED MINNESOTA PLAN

A. General

Table 3 provides an overview of the curriculum coverage provided by the Minnesota plan. Each row of this table represents a grade level; each column represents a particular school year. The numbers in each cell indicate the grade level, and the abbreviation in each cell represent the subject matter areas to be assessed. Although the subject designation is placed in distinct grade level cells, it should be noted that a small number of eligible students selected at random from other grade levels will be included in the sample to provide for age level comparisons (9, 13, and 17-year-olds) with the NAEP results. This table will be referenced and amplified in the ensuing paragraphs of this section.

Curriculum coverage is delineated in this section to the grade and age levels assessed, the subject matter areas included in the assessment program, and the reassessment cycle for measuring educational progress. Each of the components is discussed separately below.

B. Grade and Age Level Coverage

As previously mentioned, NAEP results are reported only for 9, 13, and 17-year-old students and for young adults in the 26-35 age group. The Department of Education decided not to assess out-of-school 17-year-olds and young adults in the 26-35 age group. As a result, the Minnesota statewide assessment sample will be designed to provide for adequate statistical precision to compare Minnesota statewide results for 9, 13, and in-school 17-year-olds with the results for these same age groups in the Nation and Central Region.

Educational planning in the public schools is currently based more on information grouped by grade levels than by age levels. In addition, the grade level groupings are more conducive to gathering information relative to desired outcome measures for the state; that is, it would be easier for an educator or layman to formulate the desired education performance of 8th graders, as opposed to formulating desired outcomes for 13-year-olds who may be scattered in grades 6 through 9 (the concept of

Table 3

CURRICULUM COVERAGE OF THE PROPOSED MINNESOTA TEN-YEAR ASSESSMENT PLAN

(Subject Areas[1/] by Grade and Age Levels[2/] by School Year[3/])

1972-73	1973-74	1974-75	1975-76	1976-77	1977-78	1978-79	1979-80	1980-81	1981-82	1982-83
								1	2	3
							1	2	3	4 C,A
						1	2	3	4 W,Cod	5
					1	2	3	4 Mu,S	5	6
				1	2	3	4 M,SS	5	6	7
			1	2	3	4 R,L	5	6	7	8 C,A
		1	2	3	4 C,A	5	6	7	8 W,Cod	9
			3	4 W,Cod	5	6	7	8 Mu,S	9	10
1	2		4 Mu,S	5	6	7	8 M,SS	9	10	11 C,A
2	3	4 M,SS	5	6	7	8 R,L	9	10	11 W,Cod	12
3	4 R,L	5	6	7	8 C,A	9	10	11 Mu,S	12	
4	5	6	7	8 W,Cod	9	10	11 M,SS	12		
5	6	7	8 Mu,S	9	10	11 R,L	12			
6	7	8 M,SS	9	10	11 C,A	12				
7	8 R,L	9	10	11 W,Cod	12					
8	9	10	11 Mu,S	12						
9	10	11 M,SS	12							
10	11 L	12								
11 R	12									
12										

1/ The ten subject matter areas covered in the assessment are abbreviated as follows:

R = Reading M = Math SS = Social Studies Cod = Career and Occupational Development

L = Literature S = Science Mu = Music

C = Citizenship W = Writing A = Art

2/ 9-year-olds will generally be distributed across grades 2-5; 13-year-olds across grades 5-9; and 17-year-olds across grades 9-12.

3/ Grade 8 and 13-year-old students are tested in October-November; Grade 4 and 9-year-old students are tested in January; and Grade 11 and 17-year-olds are tested in March-April.

desired outcomes is discussed in Section VII below). The statewide sample size will, therefore, be designed so that assessment results can also be reported by grade levels. As indicated in Table 4, the greatest percentages of 9, 13, and 17-year-olds in the United States are found in grades 4, 8, and 11 respectively. Since these grade levels provide the highest overlap with the NAEP age groupings, the total state sample required for reporting results by these three grade levels and by the three NAEP age levels would be smaller than for any other combination of three grade levels. In addition to providing for assessment cost economies, these grade levels do represent key stages in the elementary and secondary education ladder.

In order to increase the validity of the Minnesota and NAEP comparisons, Minnesota will follow the NAEP plan of collecting assessment data on 13-year-olds and 8th graders in the fall (October-December), on 9-year-olds and 4th graders in the winter (January-February), and on in school 17-year-olds and 11th graders in the spring (March-April). NAEP also collects data on, and reports results for, 17-year-olds who are out of school and young adults age 26-35. Minnesota, realizing the expenses involved in collecting data from these students and young adults, has given these options a low priority for the assessment program.

C. Subject Matter Areas Covered

The subject area coverage for the Minnesota assessment plan as outlined in Table 2 is initially limited to the ten NAEP subject matter areas. These ten subject areas do provide for rather comprehensive curriculum coverage. Nevertheless, the Minnesota Department of Education hopes to expand this coverage, as soon as resources permit, to include health and physical education (a measure in the psychomotor domain) and measures in the affective domain. The program is also flexible enough that additional subjects could be added, or a currently included subject area could be dropped, depending upon future events.

The decision as to which subjects are to be assessed when was greatly influenced by NAEP policy regarding the release of exercises for state and/or local use. It is desirable to use these exercises as soon as possible after they have been released in order to better insure their content "currency," as well as to reduce the time lapse between NAEP and Minnesota comparisons. NAEP exercises are usually not released before the middle of the school year following the school year in which they were used. Given the time required to screen these exercises for their relevancy to state objectives and to develop supplemental exercises, it is not feasible to plan to use the NAEP released exercises sooner than two years after they have been used in National Assessment. Hence, the Minnesota assessment schedule in Table 3 is generally two years behind the NAEP plan (refer to Table 2).

Table 4

ESTIMATED PERCENTAGE OF 9, 13, AND 17-YEAR-OLD STUDENT ENROLLED IN VARIOUS GRADES ACROSS THE UNITED STATES*

Estimated Percentage of 9-Year-Olds by Grade Level

Grade Level	Percent
2nd or lower	1.6
3rd	17.1
4th	72.5
5th or higher	8.8
Totals	100.0

Estimated Percentage of 13-Year-Olds by Grade Level

Grade Level	Percent
6th or lower	4.9
7th	18.6
8th	67.3
9th or higher	9.2
Totals	100.0

Estimated Percentage of 17-Year-Olds by Grade Level

Grade Level	Percent
9th or lower	5.8
10th	18.0
11th	59.7
12th	10.0
Not Enrolled	6.5
Totals	100.0

* Estimated from data in U.S. Bureau of the Census, Current Population Reports, Series P-20, No. 222, "School Enrollment: October 1970," U. S. Government Printing Office, Washington, D. C., June 28, 1971.

The exact NAEP groupings of subject areas for the Minnesota program have been altered somewhat based on NAEP's experience and Minnesota's preference for covering only two subject areas each year; e.g., Math and Science were separated in this plan since giving Math and Science together provided NAEP with problems due to the similarity of exercises.

Assessing only Reading (R) at grade eleven in the spring of 1972-73 stems from the desire of the Department of Education to maintain the assessment thrust or momentum that was initiated last year. Although the Department can have supplemental Reading performance items prepared for a spring assessment of 11th graders (17-year-olds), they cannot get the Literature exercises prepared in time. Doing Reading this year and only Literature at grade 11 next year does not really have an adverse effect on the long-range cycle as discussed below. This spring of 1972-73 assessment in Reading has been designated as Phase 1 of the program, the subsequent ten years of the assessment program as Phase 2.

D. Testing Cycle

By following each grade level diagonally in Table 3, it can be noted that the three grade/age levels are reassessed in the same subject area in five-year cycles--a pattern that not only allows a reasonable time for measuring the educational progress, but also provides time for the preparation of the exercise items required in the reassessment. For example, the areas of Reading and Literature (R,L) that are initially assessed in 1973-74 are reassessed in 1978-79. The Reading performances of Minnesota 9, 13, and 17-year-olds and 4th, 8th, and 11th graders in 1972-73 can thus be compared to the Reading performances of students who are in these same grade/age groupings in 1978-79.

One can also compare the performances of 4th, 8th, and 11th graders during the same year and on the same subject area. However, keep in mind that these groups are assessed at different times within the school year and such comparisons would have to be interpreted with this point in mind--especially when common exercises are given to different grade or age level groups.

V. REPORTING VARIABLES

A. General

A reporting variable is defined as a primary characteristic or set of characteristics that serves to define the group of students for which information is desired and for which output measures are to be reported. Each reporting variable has reporting groups. For example, sex is a reporting variable; males and females are reporting groups.

If all the specific information needs of Minnesota state level educational decisionmakers were known, the task of delineating reporting categories would be relatively straight-forward. Unfortunately, this is not the case. Valuable insight into these information needs was provided by feedback obtained during the course of this study, but it is apparent that all the information needs were not clearly defined and specified--a task that one could not reasonably expect to accomplish within the limited scope of this research.

Under these conditions, a feasible strategy to follow in delineating reporting categories would be to supplement the information needs that have been delineated by presenting assessment data in several general ways. Statewide problem areas can then be better defined by a wide range of users who can make certain decisions relative to attacking, exploring, or alleviating these problems. As the assessment program becomes operational and assessment results are reported to decisionmakers, feedback and followup efforts designed to determine who is making what kind of decisions on the basis of these assessment data can be used to modify the reporting categories for future assessments. As a result, this discussion of reporting categories, though it applies generally to both phases of the assessment program, is geared more toward the Phase 1 assessment tentively scheduled for the spring of 1973.

A second consideration with respect to delineating reporting variables involves the precision of the output estimates being reported. When splitting the total state sample into reporting groups, the sample size of each group must be of sufficient size to make precise estimates from the sample for each population group to be analyzed and reported. The number and nature of the reporting categories included in the program thus controls the number of students to be assessed, a factor which greatly affects data collection costs. Each additional reporting category also adds to the assessment costs by increasing the time and effort involved in questionnaire development, data analysis, report writing and reproduction, and the dissemination of results.

Since an exact determination of the amount of funds available for the Phase 1 assessment has not yet been made, it was not possible to make a final selection of the number and type of reporting categories. These categories are thus discussed in broad general terms, and their reporting groups have not been specified. Precise definitions would be made when Phase 1 becomes operational and exact funding constraints are known. The final specifications of these reporting variables must be completed, however, prior to the design of the questionnaires for collecting the supplementary student and school background information required to place students in the proper reporting category.

B. Candidate Variables

The major reporting variable is the state as a whole. It is these results which will be used in making the Minnesota versus NAEP National and Central Region comparisons.

Figure 1 contains a map of the state outlining the eleven planning regions that could be used in defining the primary reporting variables for within the state. Ten potential reporting groups are constructed by combining planning Regions 1 and 2 to form one reporting region, with the nine remaining planning regions each serving as a reporting region. This ten region group was selected over the alternatives of using each of the state's 435 school districts as a reporting group, or of further collapsing the eleven planning regions into only five reporting groups.

As discussed in detail in Chapter 5 of this report, the state sample will be selected to insure that the results reported for each of these ten regional groups have a adequate degree of statistical precision. In addition, using these ten regions as a stratification variable is a proper sampling strategy in that it also generates a state sample that is large enough to provide enough statistical precision for reporting statewide results by 20-25 other reporting variables--even though the Department of Education might decide at a later date not to use the ten planning regions as reporting groups. However, in order to maintain sufficient statistical precision, the number of reliable reporting groups within any single reporting variable, or within any combination of reporting variables, would not normally exceed eight (the true limiting factor would be the sample size for the group which has the smallest sample); e.g., one could report results by sex and four levels of SES.

The candidate reporting variables for the Minnesota assessment are listed below under three general headings: Student Variables, School or School District Variables, and Process Variables. Within each of these headings, the reporting variables are listed in a rather coarse priority sequence based on the feedback obtained from the Minnesota advisory groups (e.g., the Assessment Advisory Council and the Technical Advisory Committee). This list is not final and is subject to revision and further study as the assessment program becomes operational. In particular, the use of teacher characteristics in defining these reporting variables was subjected to considerable debate by the Assessment Advisory Council. The issue was never completely resolved, but the council indicated that a review committee would be organized to pass final judgment on these variables.

1. Student Variables
   a. Age level.
   b. Grade level.
   c. SES.

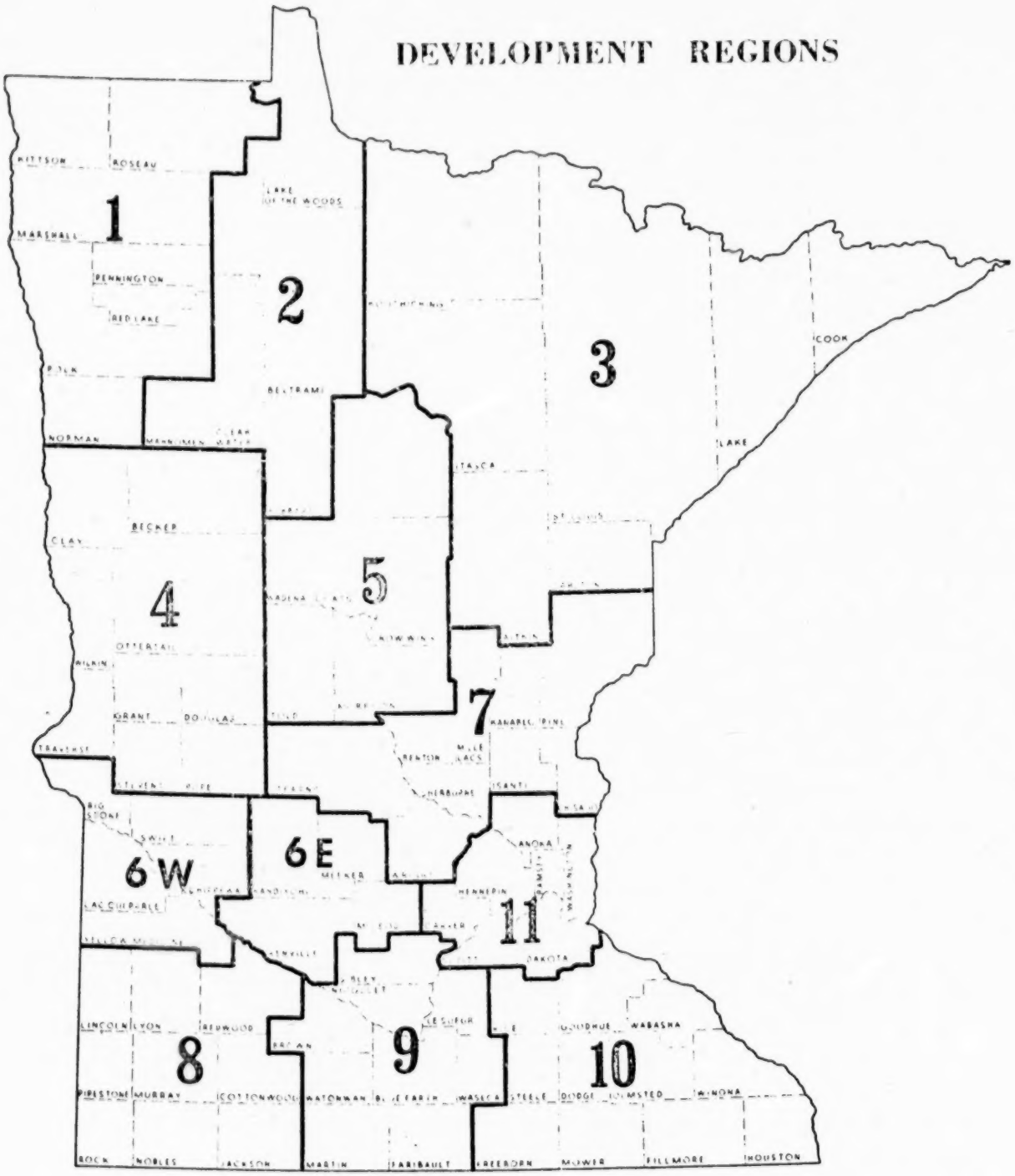


Fig. 1. Eleven Planning or Development Regions of the State of Minnesota.

d. Sex.

e. Longevity in the Minnesota Educational System.

f. Type and degree of participation in extra-curricular activities.

g. Race.

2. School or School District Variables

a. Public and/or non-public status (use three reporting groups: Public and non-public together; public; non-public).

b. Type and size of community in which school is located.

c. Principal's perception of adequacy of personnel of his school.

d. School district enrollment.

e. School enrollment.

f. Teacher/student ratio.

g. School fiscal categories.

h. Principal's perception of the adequacy of his school's resources.

i. Regional location in terms of the ten planning regions described above.

j. Racial composition of student body.

k. SES measure of student body.

l. Percent teachers with graduate degrees or with varying years of teaching experience.

3. Process Variables

a. Group students by special educational processes in the measured subject areas; e.g., code emphasis versus meaning emphasis in Reading, or traditional versus modern approach in Mathematics.

b. Group by current and previous participation (within previous 3 or 4 years) in special remedial or nonremedial programs (e.g., Title I programs).

VI. THE PIGGYBACK OPTION

The piggyback option is defined as giving local school districts the opportunity of linking up with, or piggybacking onto, the state assessment program to obtain assessment results for their districts. The Minnesota Department of Education has endorsed this option as a policy for the assessment program, but plans for implementing this option have not yet been finalized. The Department of Education has wisely decided to explore this option with one or two local school districts during the Phase 1 assessment to gain the experience of taking on the added responsibility of providing advisory assistance to local school districts wishing to pursue this option. Hopefully, procedures and policies will be developed during Phase 1 which will allow for a fuller implementation of this option at the initiation of Phase 2 of the assessment program.

Local school districts desiring to exercise this option would be responsible for any additional costs that their request might entail; e.g., obtaining assistance in designing the sample for the district, printing additional questionnaires and exercise booklets, collecting the data, analyzing the data, and reporting and disseminating the results. (The reporting and dissemination of local district results would be under the control of each district.) In other words, the entire state assessment model as described in this report could be implemented within the district. It would even be desirable for the district to supplement the NAEP/state developed exercises to reflect local education objectives.

Arguments analogous to those presented in Section III.B. relative to Minnesota following the NAEP model can be used by local school districts to justify the use of this piggyback option. Local school districts would have an added advantage of comparing their results to the statewide results. In order to make these "local district versus state comparisons" more reliable, the local district should be conducted in conjunction with the state assessment; i.e., testing the same grade/age levels in the same subject matter areas within the same calendar time frame.

VII. DESIRED OUTCOMES

The Minnesota Assessment Advisory Council has strongly recommended that the Department of Education, through the joint involvement of educators, parents, citizens, and students, pursue the development of "desired outcome" measures that would reflect desired performance levels for Minnesota students in the assessed subject matter areas. This judgmental information could be used by educational decisionmakers (includes lay people as well as professional educators and legislators) to identify areas of concern and commendation within the state's educational program.

A general approach to obtaining relevant desired outcome information could involve obtaining estimates from Minnesota citizens (parents, students, and educators) as to the percent of the state's students who should be able to answer correctly each exercise used in the assessment. These estimates could be collected at the same time that the assessment exercises are being administered to students. In addition to getting a desired outcome measure for each exercise, it might also be desirable to obtain similar estimates on groups of homogeneous exercises that can be clustered into a major theme or major skill area; e.g., to be able to state as a desired outcome that fourth graders should be able to correctly answer 10 of the 12 exercises related to the "Initial Consonant" theme in Reading "Word Attack Skills," or to state that Minnesota fourth grade students should be able to answer correctly 15 of the 20 exercises that measure Reading "Word Attack Skills."

The Technical Advisory Committee, however, posed several problems that are involved in collecting and summarizing the desired outcome data under such an approach. Some of these problems, though their solutions would be fairly time consuming and costly, could be solved by employing methodologies and procedures which are rather straightforward; e.g., design and field test questionnaires and/or interview formats to solicit from respondents, who have a wide range of competencies in specific subject matter areas, a reasonable expectation of what students should do and not an expectation of what students would do.

Other problems are not so easily solved. For example, a high school science teacher might be more biased than a housewife or French teacher with respect to the amount of science knowledge that a Minnesota eleventh grade student should have. In addition, this high school science teacher's expectations might be based on extensive experience and association with middle class, college bound students taking high school science courses; this same teacher might have had little association with eleventh graders who have not taken similar courses. Or one respondent's desires might be based on years of living in a ghetto area and doing manual labor; another's might be based on the living experiences gained in a high socioeconomic environment and on being employed in a career required a high degree of formal schooling. How does one combine the desired outcome data obtained from different groups of respondents, each of which might have a different reference base for their expectations? Should responses from all participants be weighted equally in getting desired outcomes for the state? Should the desired outcomes as expressed by different groups be presented separately? If presented separately, what would be their utility for educational decisionmaking?

Other questions as posed by the Technical Advisory Committee, concern the two basic alternatives for obtaining desired outcome information. That is, does one collect desired outcome information on individual exercises or on broad themes (as measured by several exercises)? Which has greater decisionmaking value? Which would be easier to interpret? Which has greater validity?

Despite the myriad of difficulties involved in trying to obtain desired outcome information, the Technical Advisory Committee agrees with the Assessment Advisory Council that the state assessment program should move in this direction--provided required resources could be made available at no expense to the remainder of the program. The Technical Advisory Committee thus recommends that the initial "desired outcomes" effort be limited to a small experimental, pilot approach to the undertaken in conjunction with the Phase 1 assessment.

VIII. SUMMARY

The proposed Minnesota statewide assessment program can be an important component of the state department planning function. It can provide a means of periodically monitoring achievement in the cognitive, affective, and psychomotor domains to determine wwhether or not Minnesota children know and can do those things that they should be able to do in order to live a full life.

Assessment results, however, can easily be misinterpreted--especially with respect to making "cost/effectiveness" type decisions. Assessment results generally serve only to describe the achievement status (by Reading or any other output variable) of selected groups of students at the time the measures were obtained. As such, they serve to spotlight achievement discrepancies between selected groups of students. Additional state resources (generally as a research function) can then be focused on "why" such discrepancies do in fact exist and "what" can be done about them. The Assessment results themselves DO NOT SHOW CAUSAL RELATIONSHIPS.

The ongoing National Assessment of Educational Progress (NAEP) project provides the model for the proposed Minnesota assessment program. NAEP provides for a systematic, continuous, census-like survey of knowledges, skills, understandings, and attitudes as exhibited by students and young adults in four age levels and across ten different subject areas. By following the NAEP model, Minnesota can: (1) reduce the costs of developing and scoring assessment exercises; (2) compare the assessment results for its students in the ten subject matter areas against those obtained by NAEP for students in the Nation as a whole, as well as in the Central Region of the Nation (all comparisons will be made at the same student age levels); and (3) take advantage of past and future test administration, data collection, sampling, and data analysis methodologies that have been (and will continue to be) developed by NAEP. Hence, the ten-year assessment program proposed for Minnesota would initially follow the NAEP plan of assessing the ten subject matter areas (Art, Career and Occupational Development, Citizenship, Literature, Mathematics, Music, Reading, Science, Social Studies, and Writing) in a state sample of students in three age levels (9, 13, and 17-year-olds enrolled in school). Since state educational planning is currently based more on data grouped by grade levels than by age levels, the Minnesota plan also provides for an assessment of students in grades 4, 8, and 11. Two subject matter areas are to be assessed annually at the three age and three grade levels; after five years, the same assessment cycle would be repeated in order to provide an evaluation of educational progress throughout the state. State assessment data would also be collected as per the NAEP model by trained exercise administrators, using paced tapes to standardize the administration procedures.

The utility of the assessment results depends upon how groups of students are defined for reporting purposes. The state as a whole is the major reporting variable. These state results will be compared to those of the Nation and Central Region. Ten planning regions of the state constitute the primary reporting groups for within the state. A ranked list of candidate student, school, and process reporting variables and a strategy for defining the type and number of additional reporting variables and groups has been developed.

A feature of the state assessment program that should prove valuable to local educators is an option whereby local school districts could link up with, or "piggyback" onto, the state assessment to obtain results for their districts. This option has been endorsed by the State Department of Education but will be implemented only on a pilot basis during Phase 1 of the assessment program.

The Minnesota Assessment Advisory Council has strongly recommended that the Department of Education, through the joint involvement of educators, parents, citizens, and students, pursue the development of "desired outcome" measures that would reflect desired performance levels for Minnesota students in the assessed subject matter areas. This judgmental information could be used by educational decisionmakers (includes lay people as well as professional educators and legislators) to identify areas of concern and commendation within the state educational program.

A myriad of technical problems are involved, however, in formulating the desired outcome measures for the state. Nevertheless, the Technical Advisory Committee agrees that the state assessment program should move in this direction--provided required resources are available. The Technical Advisory Committee thus recommends that the initial "desired outcomes" effort be limited to a small experimental, pilot approach to be undertaken in conjunction with the Phase 1 assessment.

Chapter 3

Management and Staffing

I. INTRODUCTION

The ten-year assessment program recommended in the previous chapter of this report involves the annual collection, analysis, and dissemination of a substantial amount of educational data. That is, with the exception of the first and last years of the assessment cycle, Department of Education personnel will be simultaneously involved in collecting data for the current year's assessment, analyzing and disseminating the results for the previous year's assessment, and preparing for next year's assessment (e.g., developing exercise items and finalizing the data collection strategy). The recommended program, by following the NAEP model of collecting data on 13-year-olds in the fall (October-December), on 9-year-olds in the winter (January-February), and on in-school 17-year-olds in the spring (March-April), does serve to reduce the data collection work load by distributing it over the school year.

An assessment program of this magnitude entails a wide range of work tasks, many of which require highly specialized technical and professional skills. This chapter presents a brief description of these required work tasks and desired staff competencies, a plan whereby staff resources could be effectively organized and managed, and a general strategy for building and developing this assessment staff within the Minnesota Department of Education.

II. MANAGEMENT AND STAFFING PLAN

A. Staffing Plan

The general work tasks involved in the annual assessment operation are listed in Table 5. A wide range of staffing plans could be designed for performing these tasks. Minnesota could, for example, build an assessment staff that would have the capability of completing all assessment tasks "in-house"--or, at the other end of the spectrum, the state could elect to minimize staffing requirements by contracting with an organization to provide a "turn-key" assessment; i.e., completing all tasks outside the Department of Education.

One could support the extreme "turn-key" approach by arguing that, instead of building a large staff of assessment technicians, the state should gear its staff requirements more toward the utilization of assessment findings to better assure that indicated program changes are designed and implemented. The ineffectiveness,

Table 5

GENERAL WORK TASKS FOR THE RECOMMENDED MINNESOTA ASSESSMENT PROGRAM
(3 AGE/GRADE LEVELS, 2 SUBJECTS, 10 REPORTING REGIONS)

Goals and Objectives

1. Determine broad goals for education in the state.
2. Define the broad state goals in operational and behavioral terms so the degree of attainment can be measured.

Information Needs

1. Delineate information needs of decisionmakers.

Instrumentation

1. Review available test exercises (NAEP, IOX, etc.) for relevance to Minnesota Objectives.
2. Prepare new exercises as required.
3. Conduct tryouts and finalize new exercises.
4. Design necessary school, teacher, and pupil questionnaires.

Data Collection (Survey Operations)

1. Package exercises and print exercise booklets.
2. Print questionnaires.
3. Hire and train field administration staff.
4. Contact schools.
5. Distribute materials.
6. Collect data.
7. Edit data.

Scoring and Data Reduction

1. Score exercises (including open-ended).
2. Reduce data on magnetic tape files for analysis.

Sampling and Data Analysis

1. Design and select sample.
2. Specify format of magnetic tape files.
3. Analyze data.

Report Preparation

1. Prepare Technical Report.
2. Prepare Highlight Reports.

(continued)

Table 5 (continued)

Dissemination
1. Design and implement pre-assessment strategy to gain public awareness, understanding, and support of assessment program.
2. Design report formats for different types of audiences.
3. Prepare news releases.
4. Design AV aides for disseminating results.
5. Disseminate results to various decisionmaking groups, special interest groups, and general public.

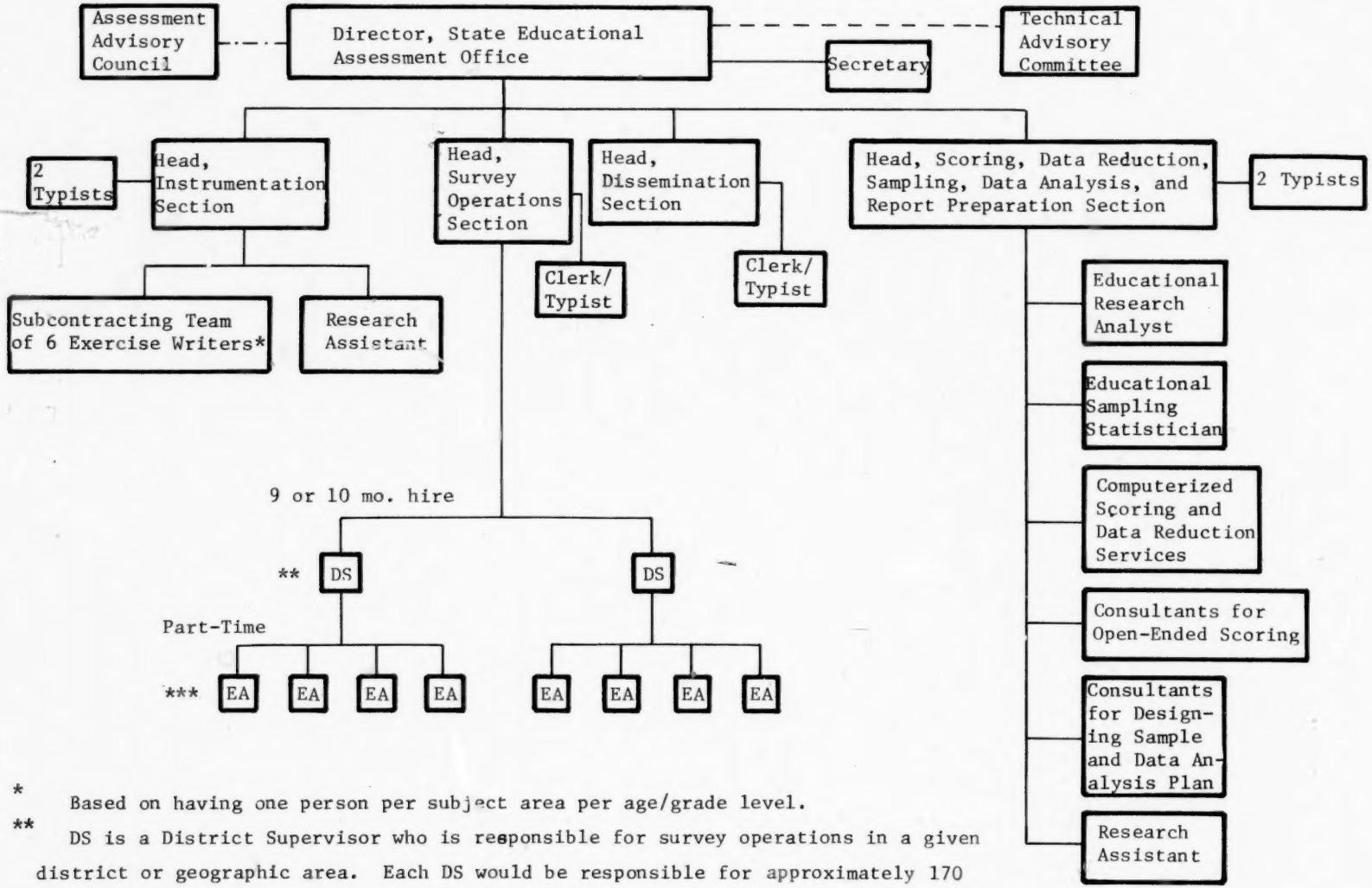
from a cost viewpoint, of hiring and retaining full time staff members who possess the variety of specialized expertise that is needed at infrequent intervals throughout the year is another effective argument against the "in-house" extreme. Arguments often used to oppose the "turn-key" extreme include the availability of funds for new staff positions as opposed to the nonavailability of contract monies, or the possibility of reducing assessment costs by developing a large and capable "in-house" assessment staff. One could pose similar arguments for the many combinations of "staffing and contracting" approaches that might lie between these two extremes. In effect, the most desirable management/staffing plan for a given state will depend to a large extent upon the state's policies, resources, geography, labor market, etc. (Minnesota, for example, is currently operating under a "freeze" policy that prohibits the hiring of additional staff.)

The Minnesota Department of Education has expressed a desire to move toward the development of an "in-house" assessment capability--even though the development of this capability might extend over a period of years. Given this charge, the management and staffing plan, as depicted in Figure 2, represents the minimal staffing requirements for the adequate completion of the work tasks listed in Table 5. It does not, however, allow for complete in-house capability in that consulting services are utilized in exercise development, scoring open-ended exercises, sampling, and data analysis. Hiring full time exercise developers and scorers for open-ended questions in ten subject matter areas, when only two subject areas are assessed per year, is just not practical. A similar argument is presented against trying to attract and retain personnel with highly specialized sampling and data analysis expertise. If the required "computerized scoring and data reduction services" are not available to the Department of Education through the Information Systems Division of the Department of Administration, it would also be necessary to contract for these services.

The functions of the advisory staff and the general duties of key personnel are briefly described below. These general descriptions, used in conjunction with the methdology discussions presented in subsequent chapters, provide a good profile of the skills and competencies required in key staff positions.

1. Director, State Educational Assessment

The director is responsible for the overall management of the program. The director, in addition to maintaining close contact with the "line" staff, must anticipate those decision points which will require technical or special inputs from the advisory staff and make sure that these inputs are made at critical decision points.



* Based on having one person per subject area per age/grade level.

** DS is a District Supervisor who is responsible for survey operations in a given district or geographic area. Each DS would be responsible for approximately 170 schools during each exercise administration period.

*** EA is an Exercise Administrator who is responsible for administering exercises within a given district.

Fig. 2. Management/Staffing Plan for Minnesota Educational Assessment Program.

It is also essential that the director stay in touch with

> ... legislators interested in education, leaders in state government who are interested in education, leaders among the local schools, and the leaders among the lay public who are interested in education. The purpose of keeping in touch is to sense the educational problems, the kinds of concerns about education, the notions that are held that may be true or false, and other things in the general climate of the environment to suggest the strategic areas for assessment, the kinds of information that will need to be provided to deal with the concerns and apprehensions, and the people to be involved in order to get support for an ongoing program. [Ref. 7]

This information, along with the progress or lack of progress being made in the ongoing program, can be used in planning modifications to the long-range program.

The director is also responsible for the technical planning of the assessment program; i.e., determining the technical procedures required to perform the assessment tasks. These technical procedures would, in turn, lead to the development of a step-by-step outline of specific jobs to be performed, as well as to suggest additional types of competencies that should be included on the advisory staffs.

2. Advisory Councils

Both of these councils, the Assessment Advisory Council (AAC) and the Technical Advisory Committee (TAC), are currently operating in a highly effective manner in Minnesota. The authors of this report, having worked closely with these councils during this planning study, were most impressed with the professional integrity, motivation, and cooperative spirit of every member of the two councils.

Representation on AAC was solicited from professional educational organizations, other state agencies, higher education, non-public education, the Governor's Office, the State Legislature, citizen groups concerned with education, and representatives from the Department of Education. (See Appendix A for a list of members of this council.) AAC, by providing for the expression of opinions and concerns of the various groups to which the Department must communicate, thus serves as a vehicle for the coordination of the assessment program.

TAC members, on the other hand, were selected for their technical expertise in areas related to educational assessment, measurement, and evaluation. (TAC members are listed in Appendix A.) The five members of this council thus provide advice on technical matters.

The responsibilities of AAC and TAC are limited to general policy establishment; specific responsibilities relative to implementation remain with the Commissioner of Education and his designees within the Department of Education. [Ref. 1]

3. Head, Instrumentation Section

The Head of the Instrumentation Section reports directly to the State Assessment Director and is responsible for the development of the instrumentation

required in the assessment program. In addition to being responsible for those tasks listed in Table 5 under "Instrumentation," this person would assume a leadership role in defining broad state goals in operational and behavioral terms so that the degree of attainment can be measured. This task would involve the development of a training program for State Department of Education staff and for school personnel in translating state objectives into measurable objectives. This person must also be adept in organizing and coordinating "working groups" of curriculum specialists and exercise writers in specific subject matter areas.

This section head would be assisted by a research assistant, a subcontracted team of exercise writers who have expertise in specifying objectives in behavioral terms, and two typists. The subject matter expertise of these exercise writers would vary with the content areas being assessed.

The scope and magnitude of the work tasks for this Instrumentation Section, plus additional comments relative to selecting and organizing the subcontracted team of exercise writers, are presented in Chapter 4.

4. Head, Survey Operations Section

The head of the Survey Operations reports directly to the Director and is in charge of all aspects of data collection for those schools selected in the assessment sample. The specific work tasks for which this section is responsible are listed in Table 5 under "Data Collection." This position requires expertise in survey operations and field interviewer management, as well as knowledge of strategies for initiating and sustaining willing cooperation among participating schools.

The head of this section will be assisted by a clerk/typist, two District Supervisors, and eight Exercise Administrators.

5. District Supervisors (DS)

These personnel have the direct responsibility of making initial contacts and all data collection arrangements with respect to sample schools in their assigned geographic area. They must communicate with school superintendents, principals, and teachers; hire, train, and supervise Exercise Administrators (see below); schedule exercise administrations within a school; select the student sample from a list of eligible students attending each sampled school; and handle routine field problems. The DS is primarily responsible for insuring that the exercises are administered properly and that discrepancies, inconsistencies, and inaccuracies are corrected before the data are sent to the scorer. It will also be necessary for the DS to assume the role of an Exercise Administrator during peak exercise administration periods.

Each DS reports directly to the Head of Survey Operations and each is assisted by four Exercise Administrators. The DS could be hired on a school-year or ten-month basis.

6. Exercise Administrators (EA)

These personnel are responsible for administering exercises to students, collecting completed questionnaires from sampled students and principals, and editing the data to assure its quality. The EA should have teaching and/or exercise administration experience. Each EA will be hired on a part-time basis and will generally be assigned to work in schools or areas near his home. EAs will be closely supervised by a DS.

7. Head, Dissemination Section

The dissemination of pre-assessment information and the dissemination of assessment results are key functions in assessment programs. The goal of this section is to put information into the hands of those who will make use of it. In addition to possessing a "public relations" personality, the head of this section should have a background in mass communications, speech, and journalism.

This person will be assisted by a clerk/typist in performing those tasks listed under "Dissemination" in Table 5.

8. Head, Scoring, Data Reduction, Sampling, Data Analysis, and Report Preparation Section

This staff member must be able to wear many hats. He is responsible for the tasks listed in Table 5 under the "Scoring and Data Reduction," "Sampling and Data Analysis," and "Report Preparation" headings. A strong background in statistics, the ability to summarize and analyze data, experience with computerized data files, and a proficiency in writing meaningful and significant reports are essential to this position. The head of this section must also be an excellent manager, capable of coordinating the efforts of consultants with special expertise in sampling, data analysis, and scoring open-ended exercises. He would also be responsible for the establishment of computerized data files (probably through a contractor).

This section would be supported by an educational research analyst, an educational sampling statistician, a research assistant, and two typists.

9. Educational Research Analyst

This position requires experience and training at the masters level in educational research or in educational tests and measurement. A background in statistics and the ability to analyze educational data and write meaningful reports are essential.

10. Educational Sampling Statistician

This person should have the equivalent of a masters degree in sampling and statistics, plus training in educational research. Using the consulting services of senior sampling statisticians and data analysts, the educational sampling statistician must be able to implement: (a) the selection of a probability sample of schools and students as per the sample design; (b) the estimation or weighting procedures

to compute estimates of population values as developed from the sample design, including adjustments for nonrespondents; and (c) the data analysis plan.

B. Management

All staff members should be initially assigned 100 percent of their time to the assessment program. However, it is conceivable that certain key line personnel could be assigned to assist other divisions within the Department in designing and conducting research studies and/or surveys after the operational aspects of the assessment program have become established--this would require at least two or three years experience with the program, providing there are no great changes in the scope of the assessment program. These assignments should be coordinated to coincide with slack periods in the assessment program; e.g., the survey operations section would have some free time during the summer months, the educational research analyst and the educational sampling statistician might be free in the fall of each year.

III. IMPLEMENTATION OF THE MANAGEMENT AND STAFFING PLAN

A. The Phase 1 Assessment

Implementation of the management and staffing plan described in the previous section would require a full time staff of nine senior professionals, two research assistants, one secretary, two clerk/typists, and three typists. In addition, eight Exercise Administrators are required on a part-time basis, each for approximately 100 man-days per year. (The funds required to provide consultants and/or contractual services for exercise development, sampling, data analysis, and computerized scoring and data reduction services would probably be equivalent to the annual salaries and overhead costs of one to one and one-half full time senior professionals.)

The current staff of the Assessment Project consists of a Director and typing assistance, plus two excellent advisory councils (AAC and TAC). Furthermore, a freeze policy on hiring new personnel within the State Department of Education is currently in effect. As a result, it would be impossible for the Department to implement the Phase 1 Assessment Program (the assessment of Reading for in-school 17-year-olds/11th graders in the spring of 1973) without contractual assistance.

It is recommended, therefore, that the Department implement the management and staffing plan as follows for the Phase 1 assessment:

1) Use the established advisory staff.

2) Have the Director also assume the role of Head, Dissemination Section.

3) Contract the responsibilities of the Instrumentation Section to a consulting team with expertise in the development of Reading exercises for 17-year-olds.

4) Contract the responsibilities of the Survey Section and the Scoring, Data Reduction, Sampling, Data Analysis, and Report Preparation Section to one or more organizations with capabilities in these areas.

B. The Phase 2 Assessment

Phase 2 of the assessment program begins in the fall of 1973 and extends through the spring of 1983. The Department could conceivably finalize the staffing plan in time to initiate this phase of the program with the full in-house assessment capability depicted in Figure 2. There are several factors, however, that should be taken into consideration before this recruitment strategy is employed.

1) Even after the current freeze on hiring is lifted, the process of obtaining approval for the required manpower slots will further delay personnel recruitment.

2) Since few states and/or large school systems have undertaken assessment programs of this magnitude, it will be difficult to find experienced personnel who possess those rather unique qualifications required to fill the four section head positions.

3) The assessment program involves the collection, analysis, and dissemination of information in three tight cycles (fall, winter, and spring) during each school year. If pre-assessment planning is not efficient and thorough, or if the surveys and data processing operations do not proceed smoothly, the system could break down. Undertaking such an endeavor with a newly organized and relatively inexperienced staff would be hazardous.

Given the above considerations and potential pitfalls, it is recommended that the Department follow a more conservative, graduated approach toward implementing the Management and Staffing Plan. That is, entire sections, or portions thereof, would be gradually staffed and "phased" into the program--it might take as long as two years to complete the staffing plan under this approach. For example, the Department could plan to:

1) Generally work closely with organizations experienced in state and/or national assessments in the recruitment and training of staff personnel ((4) below represents a specific example of this general strategy).

2) Hire the Head of the Instrumentation Section as soon as possible. This is a high priority position since instrumentation is an extremely critical component of the program.

3) Hire a Head for the Dissemination Section in the latter stages of the Phase 1 assessment. He could then work with the Director in disseminating the Phase 1 results, an experience that would better prepare him to assume the full responsibility for finalizing and implementing the dissemination plan for the Phase 2 assessments.

4) Contract out the 1973-74 school year assessment responsibilities for the Survey Operations Section and the Scoring, Data Reduction, Sampling, Data Analysis, and Report Preparation Section; however, every effort should be made to hire a Head for each of these sections during the 1973-74 school year assessment--on-site experience during key stages in the assessment would be extremely valuable. These two section heads, on the basis of this "on-the-job" experience and training, could recruit the remainder of their staffs and prepare to continue the program into the 1974-75 school year. After being on the job a while, these heads may decide to reevaluate the in-house assessment capability reflected in Figure 2 by either expanding or reducing the work done "out-of-house" by subcontractors.

5) Contract out only those 1975-76 assessment responsibilities that cannot be handled by the available in-house staff. If everything has proceeded as planned--assuming the strategy to develop a large in-house capability is not altered on the basis of experience--contractual and/or consulting services would be limited to those specified in Figure 2. Should unforeseen events occur, contracted services could be used to fill in staffing gaps and maintain program continuity.

The strategy described above represents one reasonable approach; others can be designed by reducing or increasing the milestone date for staffing completion, or by rearranging the priorities for section head recruitment (e.g., the Department may decide that hiring a head for the dissemination section is first priority). A graduated approach also provides the flexibility--while maintaining program continuity--of adjusting for such controllable factors as changes in manpower and fiscal policies, modifications in the assessment program based on operational experience, and the availability of desired personnel.

IV. SUMMARY

A Management and Staffing Plan has been designed to provide the Minnesota State Department of Education the in-house capability for completing most of the work tasks associated with the statewide assessment program. This plan does not

provide for the completion of all tasks in-house in that contracting and/or consulting services are utilized in exercise development, scoring open-ended exercises, sampling, and data analysis. Providing specialized expertise in-house is not cost effective when this expertise is required in brief, infrequent intervals throughout each assessment year.

Implementation of the Management and Staffing Plan would require a full-time staff of nine professionals, two junior professional/research assistants, one secretary, two clerk/typists, and four typists. These personnel would be initially assigned 100 percent of their time to the assessment program. However, by coordinating the assignments of the full-time staff to coincide with slack periods in the annual assessment program, these personnel could assist other Divisions within the Department in designing research studies and/or surveys after the operational aspects of the program have become established. In addition, eight Exercise Administrators are required on a part-time basis, each for approximately 100 man-days per year.

The current staff of the Assessment Project consists of a Director and typing assistance, plus two excellent advisory councils (the Assessment Advisory Council and the Technical Advisory Committee). As a result, it would be impossible to implement Phase 1 of the Assessment Program without contractual assistance.

A strategy for implementing the Management and Staffing Plan for the Phase 1 and Phase 2 assessments was developed. This strategy is based on a conservative, graduated approach whereby key personnel would be hired, given on-the-job training, and gradually phased into the program. This approach would initially require contracting out most of the work tasks; however, responsibility for these tasks would be gradually assumed by the Department of Education as key personnel are moved into the program. Complete staffing could be achieved within one or two years.

The recommended strategy provides flexibility for handling uncontrollable events that might arise, while maintaining continuity in the program. Furthermore, this approach would not place a newly organized and relatively inexperienced staff in the formidable position of having to implement an assessment program of this magnitude.

Chapter 4

Instrumentation Development

I. INTRODUCTION

A critical factor to any state assessment program is the development of necessary instrumentation--which, in the case of the proposed Minnesota program, includes assessment exercises and questionnaires for collecting background information on the schools and students included in the state sample. This chapter provides a general description of this critical assessment component, along with some general considerations relative to the time frame in which the annual assessment work tasks must be completed with the staff resources as outlined in Chapter 3.

Since exercise development for the Phase 1 Reading assessment of in-school 17-year-olds is in the final stages of completion, this chapter is directed primarily to the ten-year, Phase 2 assessment program. Questionnaire development for the Phase 1 assessment, however, will not be initiated until funding plans for the Phase 1 effort have been finalized.

A variation of the NAEP approach to exercise development is recommended for the proposed Minnesota educational assessment program; full implementation of the NAEP exercise development plan would be too expensive. A few of the key innovations of the NAEP approach are presented in this chapter; a detailed description of the NAEP plan is presented in reference 6.

The development of the desired outcome measures previously described in Section VII of Chapter 2 could be an important component of the state educational assessment program. However, because of the experimental nature of this concept, the initial research and development of necessary procedures and instrumentation pertinent to the development of desired outcome measures are not considered functions of the Instrumentation Section as per the staffing plan presented in Chapter 3; hence, this instrumentation is given limited coverage in this chapter. The maintenance and revision of desired outcomes instrumentation could, nevertheless, be assumed by the Instrumentation Section as soon as the concept is ready for implementation.

II. GENERAL TIME FRAME AND STAFFING CONSIDERATIONS

This section provides an overview of the strategy and annual time constraints for completing the instrument development tasks which are given in detail in Sections III and IV below.

NAEP exercises are released in December or January of the school year following their use in the National Assessment program. With few exceptions (refer to Tables 2 and 3 in Chapter 2), the Minnesota subject matter plan is designed to incorporate the released NAEP items into the Minnesota assessment program, beginning in the fall of the school year following the release of these items. For example, the NAEP Reading and Literature exercises that are scheduled for release in December or January of the 1976-77 school year are to be incorporated in the Minnesota 1977-78 school year assessment--keeping in mind that the exercises must be in the field for the 8th graders/13-year-olds in October, for the 4th graders/9-year-olds in January, and for 11th graders/17-year-olds in March.

Given these constraints, the exercise development year for each pair of subject matter areas extends roughly from November of one year to December of the next year. Key milestones for exercise development are shown in Figure 3 in relation to the exercise administration schedules for all three grade/age levels. By starting in late November or early December, the exercise development team has time to review and finalize the definition of state goals in measurable terms, before beginning their work on the released NAEP exercises. Completing the development of exercises for 11th graders/17-year-olds by December of the following year allows sufficient time for printing and distributing the exercise packages for the March administrations.

As indicated in the Staffing Plan (Section II.A.3., Chapter 3), the Head of the Instrumentation Section would be assisted by a research assistant, two typists, working groups of school personnel (curriculum specialists and teachers) who have been organized to assist in exercise writing, and a contracted team of six exercise writers. The team of contracted exercise writers would consist of local talent (e.g., professors and graduate students from the University of Minnesota or state colleges), supplemented by outside personnel knowledgeable about the development of NAEP exercises.

There is nothing magical or absolute about having "six" contracted exercise writers on this team. For planning purposes, it was felt that one person should be responsible for one subject area at one grade/age level--hence, six writers would be required to cover two subjects across three grade/age levels. It is <u>not</u> envisioned that these personnel work independently. In fact, it might be more appropriate for planning purposes to consider having a "three-man" exercise writing team in each subject area. Furthermore, the requirements for this contracted expertise could vary considerably year-by-year as a function of (1) the number of new exercises that must be developed to supplement the NAEP exercises (could range from "none" to "all"), and (2) the expertise and production capabilities of the organized working groups of school personnel and curriculum specialists.

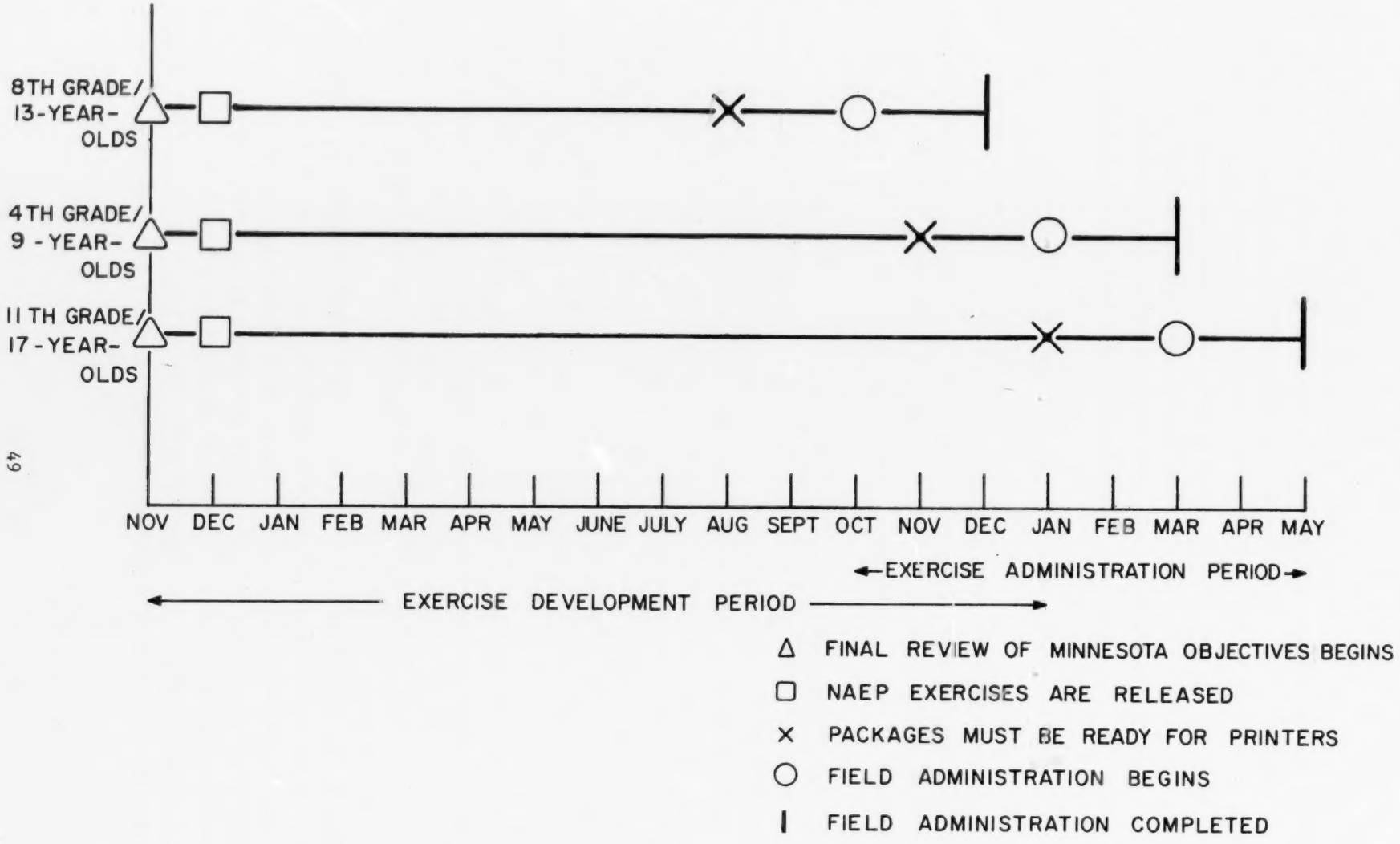


Fig. 3. Milestones for Exercise Development in Relation to Field Administration.

It is important to note that extensive supervision and control of these working groups of exercise writers is necessary to produce satisfactory results. Teachers and curriculum specialists, though experienced in developing classroom tests, do not generally have the training required for preparing state assessment exercises. Adequate provisions must be made for their orientation, training, and supervision. Mini-courses in exercise development could possibly be offered as part of their training by the Head of the Instrumentation Section and/or other measurement specialists.

III. QUESTIONNAIRE SPECIFICATIONS AND DEVELOPMENT

A. Nature and Purposes of Questionnaires

The proposed assessment program would require the development and use of three questionnaires, provided the desired outcome measures are included (refer to Section VII, Chapter 2). Each of these three types of questionnaires is briefly discussed below.

1. Student Questionnaire

These questionnaires (one for each of the three grade/age levels) are to be included in the appropriate exercise package to provide for the collection of background information on each student included in the sample. These background data will be used to group students as per the reporting variables previously discussed in Chapter 2 (Section V); e.g., age, grade level, sex, race, size and/or type of community in which the student lives, parent occupation, level of parent education, kinds of reading materials in the home, number of years educated in Minnesota school system, school programs enrolled in, and extra curricular activities participated in.

Though initially limited to the collection of student background data, these questionnaires could be expanded as time and funds permit to include items on such affective variables as student attitudes toward school, student attitude toward himself as a learner, student perceptions of the school and its programs, and student expectations regarding education.

These background factors and dimensions in the affective domain are common to a number of surveys and assessment programs, and their relationships to achievement have generally been well researched. However, their inclusion could provide a degree of validation of their effects in Minnesota, as well as provide some degree of comparability with other studies. A further and significant benefit would be possible descriptive results showing how students with these different background and affective attributes were distributed in Minnesota and some of their potential interactions.

2. School Questionnaire

This questionnaire would be designed to collect background information on each school in the sample and would be filled out by the school principal. School questionnaires would be used to collect information on such school factors as: public or non-public status; enrollment data; per-pupil expenditure data; average teacher salaries; regional location; racial composition of student body; type of community in which school is located; school staff background data; and the principal's perceptions of the adequacy of his school's facilities and staff.

3. Desired Outcome Questionnaires

If the desired outcomes concept is included in Phase 1 of the assessment program, the type and number of questionnaires required to collect relevant information will depend upon the nature and magnitude of the pilot approach to be undertaken. Pending such a decision, the brief discussion presented in Chapter 2 (Section VII) of the technical problems involved in implementing this component should suffice to provide a general description of the information to be collected by these instruments.

B. Questionnaire Development

Four major steps are involved in the development of the three questionnaires discussed above:

1) Delineate the variables and data elements to be included in the questionnaires.
2) Review available questionnaires for items that can be incorporated into the required questionnaires.
3) Develop draft questionnaires.
4) Field test and revise the draft questionnaires.

Student and school questionnaires can be prepared by the Instrumentation Section as described in Chapter 2. The initial development of the desired outcome questionnaires, however, is a potentially large and difficult task, the completion of which is beyond the scope of the resources planned for the Section. The ensuing discussion is thus directed only to the development of the student and school questionnaires.

The first task, the delineation of the questionnaire's data elements, is dependent upon a final selection of reporting variables and a definition of the reporting groups for each selected variable. Given these reporting variables and groups, the information to be collected via the questionnaire will further depend upon the availability of required information through Department of Education records. Although some of the required school information will be available through this source, the odds are small that Department of Education files would contain retrievable information specific to individual students.

Having determined the information to be gathered, questionnaires used by Federal Agencies and by other states and local districts to collect similar data should be reviewed for items which are applicable to the Minnesota program, with little or no modification. Redeveloping questionnaire items that have already been subjected to field testing and reviewing is not cost effective.

In completing the third task, the questionnaires should be designed:

1) So that the student questionnaire would take no longer than 15 minutes to complete, the school questionnaire no longer than 30-45 minutes.
2) To betransformed into computerized form by optical scanning methods that eliminate the human element as much as possible from the data transformation process.
3) To be compatible with the data collection procedures outlined in Chapter 6.
4) To expedite the data analysis plan discussed in Chapter 7.

As a final task, the draft questionnaires should be field tested on a small sample of students and principals to obtain timing estimates and to elicit critical appraisal with respect to interpreting the items and/or selecting item alternatives. The questionnaires should be revised as per the results of these limited field tests. If substantial revisions are required, the field tests should be repeated.

The reliability and validity of questionnaire responses relative to this type of educational data have been the subject of numerous research studies. In addition, similar questionnaires have already been used in large scale surveys following similar data collection and analysis plans; therefore, it does not appear necessary to invest limited resources in conducting elaborate pretests of these questionnaires to justify the sample size, to finalize the field procedures, to test scoring and data analysis procedures, and/or to investigate the reliability and validity of item responses. (Note that because of the experimental nature of the desired outcomes component, questionnaires designed to gather desired outcomes data should be subjected to pretesting.)

These questionnaires as initially designed for the Phase 1 assessment will probably require slight annual revisions based on feedback received, changes made in reporting variables, and the inclusion of some items that might be specific to the subject matter areas being assessed.

IV. EXERCISE DEVELOPMENT

A. General

The general exercise development tasks and the sequence of their performance are as follows:

1) Define the broad state goals and translate them into measurable, operational and behavioral objectives.

2) Review the NAEP objectives and released exercises and select those that are relevant to the Minnesota objectives.

3) Identify the gaps in those Minnesota objectives which are not adequately measured by the selected NAEP exercises.

4) Develop exercises to fill in the identified gaps.

5) Package the exercises.

Important considerations in the performance of each of these tasks are discussed in separate sections below.

B. Definition of Behavioral Objectives (Task 1)

Objectives written in behavioral terms are necessary to the production of good assessment exercises. The Head of the Instrumentation Section, in assuming the responsibility for completing this task, should seek a variety of inputs. That is, literature should be reviewed to learn about the current thinking of curriculum specialists; objectives developed by local districts (both within and outside of Minnesota), by other state departments, and by various other educational research organizations (e.g., the Instructional Objectives Exchange at the University of California at Los Angeles) should be examined; and working groups of curriculum specialists and teachers should be organized, perhaps on a released time basis, to assist in this task.

After a set of behavioral objectives has been drafted, they should be subjected to review by subject matter specialists, educators, and lay people. The Assessment Advisory Council would provide an excellent vehicle for the review by educators and lay people. A revised set of objectives should be developed on the basis of these reviews. After these reformulated objectives are once again reviewed, they can be finalized.

Completion of this task is a prerequisite to exercise development. However, it is expected that the "exercise development" team (i.e., the contracted exercise writers and assembled working groups of school personnel) will be required to "polish up" these objectives prior to beginning their work on Task 2.

C. Review of NAEP Materials (Task 2)

Given the lists of behavioral objectives for Minnesota and for NAEP, those released NAEP exercises which are best suited to the needs of Minnesota must be selected. This review process could be conducted by groups of subject matter specialists only, groups of lay persons only, and groups containing both subject matter specialists and lay persons. The results of these reviews would then be presented to the staff of the Instrumentation Section.

The initial step in this review process is to screen out those exercises developed to measure NAEP objectives which are not compatible to Minnesota objectives. After this step has been completed, each exercise should be screened in greater detail. Some of the key factors to be considered in further reviewing these exercise items are presented below. [Ref. 6]

1. Offensiveness

All potential exercises should be reviewed for their potential offensiveness to the Minnesota public. The involvement of lay people in this process is essential. Since all NAEP materials are subjected to a similar review, one would not expect to find anything among them that would be offensive to Minnesotans. However, the NAEP criteria for "inoffensiveness," as covered in pages 42-46 of Reference 6, should be reviewed.

2. Ease of Scoring

Some NAEP exercises require open-ended responses that are more expensive to score than the multiple choice responses which may be machine scored. Some subject areas, e.g., citizenship and writing, are difficult to measure without extensive use of open-ended or essay type exercises. In other areas some flexibility might be available for limiting the number of open-ended exercises.

3. Ease of Administration

Some NAEP items are designed to be administered to individuals; most are designed for group administrations. Minnesota would probably want to either exclude the "individual" exercises or adapt them for group administration because of the additional costs involved in individualized administration. (Caution should be exercised when interpreting the results for any NAEP versus Minnesota comparisons that are made on those exercises which were individually administered by NAEP, but were modified by Minnesota and administered to groups of Minnesota students.)

Some of the NAEP group exercises might also be excluded because of the equipment and materials required in their administration.

4. Content Validity

The content of each NAEP exercise should be examined in terms of whether or not it is assessing something important and desirable for Minnesota children to know, and whether or not it is measuring the objectives for which it was intended.

5. Content Appropriateness

Each exercise should be examined for its appropriateness for the age and grade levels for which it is being considered. For example, is the exercise stated so that the student will understand what he is to do? Is the vocabulary appropriate?

D. Identification of Gaps (Task 3)

A comparison of the results for Tasks 1 and 2 serves to identify the "exercise" gaps to be filled. However, this task is not quite that straightforward in that the size of this gap must be weighed against the availability of resources. If the gaps cannot be filled with the resources available, the results of Tasks 1 and 2 should be reevaluated in order to establish priorities for Task 4. This review could also serve to identify for further considerations those NAEP exercises that were excluded on the basis of "borderline" decisions; i.e., some of these exercises could possibly be used to plug up the gaps that cannot be filled with existing resources.

The review panels and groups involved in Task 2 should also be involved in this task.

E. Development of Exercises (Task 4)

This critical task consists of five essential components:

1) Development of prototype exercises.
2) Preparation of exercises.
3) Review and revision of exercises.
4) Field testing and revision of exercises.
   Final reviews and selections.

A detailed description of the revised NAEP plan for completing the subtasks associated with each of the above components is presented in Reference 6 (pp. 111-129). This revised NAEP approach to exercise development was based on five years of experience and is recommended for the Minnesota educational assessment program--with one important variation. NAEP contracts most of the creating, field testing, and revising of assessment exercises to outside organizations. This is a costly approach. As outlined in the management and staff plan in Chapter 3 and further discussed in Section II above, Minnesota would supplement the "contracted" local university expertise with an appropriate force of school personnel and curriculum specialists. The recommended approach, though usually less expensive, is more difficult to supervise and control to ensure satisfactory results.

Three of the key features of the current NAEP plan are especially noteworthy.

The first is the involvement of student groups to review both the objectives developed in Task 1 and the exercises developed in this task.

The second is an increased emphasis on the preparation of detailed prototype exercises and scoring keys. Each prototype exercise, since it serves as a model for the development of the other exercises, must be a detailed, complete, and concrete example of an exericse. It must be clearly stated and completely classified as to objective, grade and/or age level difficulty, time estimates, etc. In addition,

it should be followed by instructions to the administrator, a rationale for the scorer, directions to the scorer, specific acceptable and unacceptable responses, and a scheme for reporting the results. [Ref. 6, p. 116]

The third is the organization of an Exercise Development Advisory Group comprised of measurement specialists and educators. This group reviews and evaluates the prototype exercises for content validity, appropriateness, relevance, and scorability. They also participate in the final reviews and selection of exercises from the pool of field exercises.

F. Packaging the Exercises (Task 5)

Each exercise package consists of a set of exercises and a student questionnaire. Approximately two hours worth of Reading exercises are planned for the Phase 1 assessment. As a result, the exercise packages will be identical in that all the Reading exercises will be administered to all students. However, when exercises for two subject areas are administered during Phase 2 assessments, it will be necessary to distribute the total number of exercises over three or four exercise packages to reduce the exercise administration time required per student. Each package for the same grade/age level will contain identical student questionnaires; however, the exercise components will vary. A form of matrix sampling will be used whereby each exercise package will contain a portion of the total number of exercises required to assess both subject areas. One exercise package would be assigned to each student in the sample on a probability basis.

An overriding consideration in packaging these exercises is to ensure compatibility with NAEP materials and procedures. Other considerations include the usual ones of exercise format, placement location, mode of administration, and procedures to ensure standardization among testing situations and conditions. With regard to placement locations, the exercises should be carefully placed in the package in a quasi-random order to avoid problems of inter-item responses. A paced-tape mode of administration will be used to ensure standardization and to minimize the effects of slow reading ability among students.

Additional considerations are required for the Phase 2 program because the total "exercise set" would be distributed over three or four packages. (The exact number of packages will be a function of the total time required to complete all the selected exercises in both subject areas and the total time that each student will be available for testing.) That is, exercises should be distributed across the packages such that the subject area coverage, administration time, and the degree of difficulty of the exercises in all packages are approximately equal. In addition, the strategy used to assign the exercises to packages must also be consistent with the data analysis plan.

For example, if results are to be presented by themes or groups of homogeneous exercises, it would be advantageous to package together all items for a particular theme.

If time and funds permit, tryouts for the exercise packages should be conducted on a small sample of students. These tryouts provide an important check on timing and other general administration procedures.

As a final comment, the package assembly should be accomplished in close cooperation with the agency or subcontractor(s) who would print and score the exercise booklets.

V. SUMMARY

Instrumentation for the proposed Minnesota state educational assessment program includes exercises for measuring performance in subject areas and questionnaires for obtaining student and school background information. If the desired outcomes measure is also included as a component of the assessment program, additional questionnaires would be included in the instrumentation to be developed. Because of the experimental nature of this desired outcomes component, the Instrumentation Section would require additional resources for the initial development and pre-testing of the required questionnaires.

Exercise development for each school year assessment can generally begin with the release of NAEP exercises in December or January of the preceding school year. There are exceptions in the proposed Minnesota ten-year program where exercise development can begin sooner because of a greater time lag between the release date of NAEP exercises and their scheduled incorporation in the Minnesota program. Development of exercises for the annual spring assessments of in-school 17-year-olds should be completed in December, approximately two months prior to their use. This amount of "lead time" is required for packaging, printing, and distributing assessment materials in time for the scheduled field operations.

The magnitude of the annual exercise development task would vary depending upon the degree to which Minnesota objectives are measured by the released NAEP exercises. The composition of the exercise development team, which would generally begin its 12-month-task in December, would also vary annually as different subject matter areas are rotated into the program. The Head of the Instrumentation Section, and his in-house exercise writing group, would be complemented as required by outside consultant "exercise writing" expertise in the appropriate subject areas and grade/age levels.

Major steps in the questionnaire development plan include a delineation of the variables and data elements to be included in the questionnaires, a review of available questionnaires for items that can be incorporated into the required questionnaires,

the development of draft questionnaires, a field test(s) to tryout the questionnaires, and finally a revision of the draft questionnaires. As special considerations, all questionnaires should be designed to be brief, amenable to transformation to a computerized file by optical scanners, and compatible with the data collection and analysis plans.

The plan for developing the exercises required to supplement the released NAEP exercises involves the following subtasks:

1) Definition of broad state goals and their translation into measurable, operational and behavioral terms.
2) Review of NAEP objectives and released exercises in order to select those exercises which are relevant to Minnesota objectives.
3) Identification of the gaps for those Minnesota objectives which are not adequately measured by the selected NAEP exercises.
4) Development of required exercises.
5) Packaging of the exercises into booklets.

An important consideration in completing these subtasks is the involvement of subject specialists, educators, lay people, and, perhaps, students in reviewing the objectives and exercises at key stages in their development. Reviews of exercises should take into account their potential offensiveness, ease of scoring, ease of administration, content validity, and content appropriateness. Using a form of matrix sampling, these exercises could be distributed over several exercise packages to minimize the exercise administration time per student; i.e., each student in the sample could take only a portion of the total number of exercises involved in the assessment.

NAEP, on the basis of five years of experience in developing their exercise items, has formulated a revised approach to this critical task. A modification of this revised NAEP plan is recommended for the proposed Minnesota educational assessment. NAEP contracts most of the exercise development to outside organizations. It would be less expensive for Minnesota to organize, train, and supervise working groups of school personnel to assist with exercise development. These personnel would be assisted and supplemented by local consultants from the University of Minnesota and local colleges and by consultants knowledgeable about the development of NAEP exercises. Hence, the recommended plan would generally follow NAEP procedures, but would require less outside contractual assistance--especially after the program has been in operation for a few years.

Chapter 5

Sample Design

I. INTRODUCTION

The primary concern in designing a sample for the proposed Minnesota Educational Assessment program is to have the sample design produce data compatible with the analysis plan of the assessment program. Sampling, which provides educational decisionmakers and those interested in education with results of sufficient precision at a reasonable cost, was selected over the alternative of providing educational performance results based on a census of every student in the grade/age levels to be assessed.

Several alternative sample designs were considered for the assessment program. The alternatives differ in the levels of reportage of assessment results. The alternatives ranged from that producing results for the State of Minnesota only to individually reporting results for all of Minnesota's 435 school districts. After considerable review of the amount of resources projected to be available for the assessment program, the alternative of reporting results for ten reporting regions was selected for implementation.

In this chapter, the general principles one must consider for designing a sample for a statewide educational assessment and the general requirements of the sample design are first discussed. Alternative sample designs for Phase 1 (spring 1973) of the Minnesota Assessment Program are next analyzed and a recommended sample design for Phase 1 is discussed. A similar sample design will be applicable for subsequent phases of the assessment program.

II. GENERAL SAMPLE DESIGN CONSIDERATIONS, PRINCIPLES AND REQUIREMENTS

A. Principle for Stratifying Schools

Ideally, from a statistical viewpoint, the best stratification criteria to use for grouping the schools of the assessment population before sampling would be to form groupings of the schools whose educational performance in the subject areas being assessed are as identical as possible. For Reading assessment, the ideal way to stratify the schools of the assessment population would be to place schools of like Reading achievement into the same group or stratum and then select a sample of schools from each homogeneous group. The number of groups (i.e., strata) and the number of

schools in each group, as well as the method of selecting the sample schools from each group, is a function of the sample design. In practice, such educational performance data is not readily available for all schools of the state. Consequently, other criteria or data sources must be used to stratify the schools of the population being assessed. However, the data used for stratification should be correlated with achievement in order to be effective from a statistical viewpoint.

B. The Distinction Between A Stratification and A Reporting Variable

From a sample design viewpoint, there are two types of variables in an educational assessment survey; i.e., stratification variables and reporting variables. Stratification variables are those variables used to subdivide (stratify) the population before sampling so as to provide relatively homogenous groupings which, when incorporated into the sample design, give definite payoffs in the statistical precision of the assessment results. The second kind, a reporting variable, is one which is to be used in the analysis plan of the assessment. It is possible for a stratification variable to be used both as a stratification variable for sample design purposes and as a reporting variable for analysis purposes. An example of this situation in the proposed data analysis plan for the Minnesota assessment is the grouping of schools in the state into ten geographical regions for stratification purposes and then reporting results by these same ten regions.

It is also possible for a variable to be used in stratification but not as a reporting variable, and vice versa. For example, if it is decided that educational performance results should not be reported by the ten reporting regions, region could be used as a stratification variable in the sample design, and not as a reporting variable. Independent of the statistical gains derived from stratification, stratifying the schools by region before sampling serves as a guarantee that the sample will be geographically spread across the state, thus giving the sample "representative credibility."

An example of a reporting variable that is not a stratification variable in the proposed analysis plan of the Minnesota assessment is the level of high school education attained by the respondent's parents. These data are not obtained beforehand on all children and stratification on this variable is not possible; therefore, the children must be classified into these categories for reporting after the sample has been selected and assessed.

C. The Stages of Sampling and Principles for Sample Allocation

The sample designs to be considered for the assessment involve two sampling stages. Stage one of the sample design will consist of selecting a random sample of schools within each school stratum. Stage two of the selection process will consist of selecting a random sample of pupils within the schools selected at stage one. In

such a design it is usually more precise to allocate the total sample of pupils to as many schools as possible in the population being assessed. In considering costs, it is usually cheaper to confine the total number of sample students to only a small number of sample schools. Formulas for making efficient allocations, considering both precision and cost, are now available in almost any textbook on sampling theory. If the following factors are either known or very well estimated, the computations are simple and straightforward: (1) the amount of variability from one school to another with respect to the achievement area being measured, (2) the variability from one pupil to another within those schools, and (3) the relative costs of including additional schools in the sample as compared with increasing the numbers of pupils assessed within a school.

One difficulty in applying these tools to the practical problem of designing a sample for an educational assessment is that the numerical values of the required measures of sampling variability and the costs that enter into the formulas are not known with any high degree of exactitude. Fortunately, these allocation parameters can usually be approximated closely enough to allow a near optimum degree of statistical precision. Other statewide educational assessments may provide useful data in making these approximations.

D. The Concept of a "Design Effect"

The concept of a "design effect" is most important in the evaluation of the expected statistical precision of the alternative sample designs. For a two-stage sample design, the standard error of a p-value* is increased over that expected from simple random sampling by virtue of the "clustering effect" built into the sample design. That is, pupils within the same school (i.e., clustered) usually respond or perform more nearly alike than students in different schools. However, on the other hand, the stratification process usually increases precision. If the allocation of the sample to the strata is so far from optimum that large weights are required for unbiased estimates, sampling variability may also be increased. The net result of all the factors (e.g., clustering, weighting, and stratification) is measured by a design effect index which is defined as the ratio of the sampling variance of the p-value for the sample design actually used (e.g., two-stage sampling) to the variance of the p-value that would be obtained from simple random sampling. This can be expressed as:

---

* A p-value is a statistic estimating the proportion of students who respond correctly to an assessment exercise. The standard error of a p-value is an estimate of the variability of the sample p-value in repeated sampling with a fixed sample size and sample design.

$$\text{Design Effect} = \frac{\text{Sampling Variance of p-Value for Two-Stage Sample Design}}{\text{Sampling Variance of p-Value for Simple Random Sampling}}$$

Design effect values are usually greater than 1.00 indicating that the sample design actually used produces less statistical precision than simple random sampling. However, when the statistical precision of the sample design is considered relative to its cost, it is possible for a sample design with a design effect greater than 1.00 to be cost-effective. Furthermore, simple random sampling for a statewide assessment would require a list of all pupils in Minnesota from which to select the sample. This is clearly not feasible from both operational and cost considerations. The design effect values will be expected to vary from exercise to exercise. For planning purposes, a design effect of 1.15 will be used in calculating expected standard errors for the sample designs considered. This value is based upon RTI's experience in using a two-stage sample design in the Maine Assessment of Educational Progress.

E. Requirements of the Sample Design

In general, the requirements of the sample design are:

1) The sample should be a probability sample. That is, each student in a given age class or grade in a public or non-public school in the State of Minnesota should have a known positive chance of inclusion in the sample.
2) Each of the ten geographical reporting regions of the state should be represented in the sample so that results can be reported for them with nearly equal statistical precision.
3) Each reporting variable or any combination of two reporting variables can be reported for up to eight reporting groups. This requirement will allow for the analysis of interaction effects of certain pairs of reporting variables. That is, differences in the effects of one reporting variable for different levels of the remaining variable will be examined.
4) If for a given age and/or grade level the amount of time it takes to administer the entire set of assessment exercises to a given student is longer than desirable, a matrix sampling approach will be developed to shorten the length of time each pupil will be tested.
5) The school and pupil school sample sizes will be such that estimates of the sampling variability of the reported results can be estimated from the sample data.

III. PHASE 1 SAMPLE DESIGN

A. Introduction

The sample designs discussed will be those for the 17-year-old and eleventh grade populations for the school year 1972-73 since it is planned that this population be assessed on Reading Achievement in the spring of 1973. The sample design is easily generalized to other age group/grade level populations. The administration of one package of assessment exercises is assumed. The recommended sample design can be easily extended to allow for matrix sampling of additional package(s) of assessment exercises.

B. Stratification of the School Population

The requirement that results be reportable by each of the ten reporting regions is a sample design specification that will be met through stratification. Figure 1 of Chapter 2 gives a map of the State of Minnesota which geographically defines the ten reporting regions. Notice that planning regions one and two have been pooled together to form assessment reporting region one. For the 17-year-old population and for the 1972-73 school year, Table 6 gives a description of the population for the ten reporting regions.

The first step in the stratification process will involve grouping the schools into ten groups using the reporting region in which the school is located as a criterion. The next step of the school stratification process will involve grouping the schools within each of the reporting regions by school level variables which are expected to relate to educational performance. Some candidate stratification variables that will be considered for this purpose and that will be used either individually or in an index are the following:

1) State Income Tax Return Data by School District.
2) Per Pupil Expenditure by School District.
3) Size of School.
4) Size of Community in which School is Located.
5) Type of Community in which School is Located.
6) Variables available from the 1970 Census Tabulated by School District.

C. Pupil and School Sample Allocations and Their Corresponding Statistical Precisions for the Reporting Regions

Two sample allocations across the ten reporting regions will be considered. The first allocation consists of allocating the sample students to the regions in proportion to the total number of students in that grade/age in the region. That is, smaller regions (fewer students) will have smaller samples and larger regions will have larger samples. Column 3 of Table 7 gives the proportional allocation

Table 6

STRUCTURE OF THE 17-YEAR-OLD POPULATION FOR THE TEN REPORTING REGIONS

Reporting Regions	Planning Regions 1/	Number 2/ of 17-Year Olds	Number of School 3/ Districts	Number of Secondary Schools 2/	Number of 17-Year Olds Per District	Number of 17-Year Olds Per School
I	1,2	3220	51	52	63	62
II	3	7241	36	48	201	151
III	4	4041	41	41	99	99
IV	5	2594	26	26	100	100
V	6	3479	46	46	76	76
VI	7	5160	42	43	123	120
VII	8	3239	44	44	74	74
VIII	9	4402	46	46	96	96
IX	10	7917	54	57	147	139
X	11	34238	49	79	699	433
Total		75531	435	482	174	157

1/ See Figure 1 for map showing definition of planning regions.

2/ The source of the 17-year-old count data was published by State of Minnesota Planning Agency Office of Local and Urban Affairs, titled "Age and Sex of Minnesota Population." The source of their data was 1970 U.S. Census. This data was tabulated by State, the Eleven Planning Regions, and the 87 counties of the State.

3/ The data for this column was tabulated from a computer printout supplied to RTI by the Minnesota State Department of Education dated September 28, 1972, listing all the public schools in the State.

Table 7

ALTERNATIVE SAMPLE ALLOCATIONS AND THEIR STATISTICAL PRECISION
FOR THE STATE OF MINNESOTA AND THE TEN REPORTING REGIONS

Reporting Regions (1)	Number of 17-Year-Olds in Population (2)	Proportional Sample Allocation			Disproportional Sample Allocation		
		Number of 17-Year-Olds in the Sample (3)	Proportion of 17-Year-Olds in the Sample for Reporting Regions (4)	Expected Standard Error of P-Value (5)	Number of 17-Year-Olds in the Sample (6)	Proportion of 17-Year-Olds in the Sample for Reporting Regions (7)	Expected Standard Error of P-Value (8)
I	3,220	218	.0675	.0351	410	.1273	.0247
II	7,241	489	.0675	.0234	410	.0566	.0257
III	4,041	273	.0675	.0313	410	.1015	.0251
IV	2,594	175	.0675	.0391	410	.1581	.0243
V	3,479	235	.0675	.0338	410	.1178	.0249
VI	5,160	248	.0675	.0278	410	.0795	.0254
VII	3,239	219	.0675	.0350	410	.1266	.0247
VIII	4,402	297	.0675	.0300	410	.0931	.0252
IX	7,917	534	.0675	.0224	410	.0518	.0258
X	34,238	2,312	.0675	.0108	1,410	.0412	.0140
State	75,531	5,100	.0675	.00725	5,100	.0675	.00805

of the sample to the reporting regions. The formula to determine the proportional allocation for each of the regional pupil sample sizes of column 3 are:

$$\begin{pmatrix}\text{Number of 17-year-olds in Reporting Region in the Sample}\end{pmatrix} = \left(\frac{5100}{75531}\right) \times \begin{pmatrix}\text{Number of 17-year-olds in Reporting Region Population}\end{pmatrix}$$

where 5100 is the total sample size and 75,531 is the total number of 17-year-olds in Minnesota as given in Table 6.

Column 5 of Table 7 gives the standard error of a p-value* assuming a constant p-value and "design effect" from reporting region to reporting region.

Observe that the expected standard errors in column 5 range from .0108 to .0391 while the expected standard error of a p-value for the state estimate is .00725.

It can be seen that allocating the sample proportionally to the ten reporting regions does not meet the sample design requirement that results of nearly equal statistical precision be reported for each of the ten regions. To meet this need, an alternative to the proportional allocation, called disproportional allocation, is needed. The effects of disproportionate sampling on precision are shown in Table 7.

---

* The formula for the values of the standard errors of column 5 of Table 7 is:

$$\text{Standard Error} = \sqrt{\begin{pmatrix}\text{Design Effect}\end{pmatrix} \times \begin{pmatrix}\text{Formula for Sampling Variance of p-value for Simple Random Sampling}\end{pmatrix}}$$

$$\text{Standard Error} = \sqrt{\begin{pmatrix}\text{Design Effect}\end{pmatrix} \times 1 - \begin{pmatrix}\text{Proportion of Population in Sample}\end{pmatrix} \times \begin{pmatrix}\frac{(\text{p-value})(1.0-\text{p-value})}{\text{Reporting Group Sample Size}}\end{pmatrix}}$$

Using a p-value of .50, a design effect of 1.15 and a proportion allocation of .0675 we have

$$\text{Column 5} = \sqrt{(1.15) \times (1 - \text{Column 4}) \times \frac{.50 \times .50}{\text{Column 3}}}$$

$$\text{Column 5} = \sqrt{(1.15) \times (1 - .0675) \times \frac{.50 \times .50}{\text{Column 3}}}$$

$$\text{Column 5} = \sqrt{\frac{.26809}{(\text{Col. 3})}}$$

For region I, the standard error would be $\sqrt{\frac{.26809}{218}} = \sqrt{.00125} = .0351$.

The formula to calculate the standard errors of column 8 of Table 7 is the same as that given in the Equation in the footnote on the previous page except for the value of the standard error of the p-value for the state (the number in the last row) which requires a modification due to the varying proportions of the sample in the population (note that the p-value and "design effect" are assumed to remain constant from reporting region to reporting region). The major characteristics of this allocation relative to the proportional allocation is that the range of the expected standard errors over the ten reporting regions has reduced considerably from [.0108 to .0391] for proportional allocation to [.0140 to .0258] for the disproportional allocation. Under the disproportional allocation, we expect the standard error of all regional results not to exceed .0258 compared to .0391 for proportional allocation.

In addition, the standard errors for each reporting region for disproportionate allocation are in most cases smaller than those for proportionate allocation. There is a tradeoff, however, in that as the regional estimates become more precise, the overall state estimate becomes less precise.

In the present situation, the p-value for the state estimate increases from .00725 to .00805, an increase of 11 percent. The loss in precision in absolute magnitude .00080 (.00805 minus .00725) is quite small and is not alarming from a practical viewpoint. Consequently, it is recommended that the disproportional allocation be used in the Minnesota Educational Assessment sample design.

As discussed in Section II, the important factors in determining the number of pupils to be selected per school are:

1) Variability of p-values from school to school.
2) Variability of exercise responses between pupils within school.
3) Cost of including additional schools in sample.
4) Cost of including additional pupils in sample within selected schools.

Based upon estimates of these variances and cost components from other educational assessment surveys, a sample of 20 randomly selected students per school is recommended.

Once the total pupil sample size, its allocation to the reporting regions, and the number of sample pupils per school are decided upon, then the school sample sizes for each of the reporting regions is a straightforward calculation. Using the pupil sample sizes of the disproportional allocation of column 6 of Table 7 and allocating two "special" schools* to each reporting region, column 5 of Table 8 gives the total school sample sizes by reporting regions. Column 8 of Table 8, gives the largest

---

* See footnote 1 of Table 6 for a definition of a "special" school.

Table 8

REPORTING REGION SCHOOL SAMPLE SIZES AND NUMBER OF SCHOOL STRATA FOR THE DISPROPORTIONAL ALLOCATION AND THE 17-YEAR-OLD POPULATION

Reporting Regions (1)	Number of Secondary Schools (2)	Number of Sample Schools			Number of Strata		
		Special[1/] (3)	Regular (4)	Total (5)	Special (6)	Regular (7)	Total (8)
I	52	2	20	22	1	10	11
II	48	2	20	22	1	10	11
III	41	2	20	22	1	10	11
IV	26	2	20	22	1	10	11
V	46	2	20	22	1	10	11
VI	43	2	20	22	1	10	11
VII	44	2	20	22	1	10	11
VIII	46	2	20	22	1	10	11
IX	57	2	20	22	1	10	11
X	79	2	70	72	1	35	36
	482	20	250	270	10	125	135

[1/] A "special" school is one with a small 17-year-old enrollment not in grades 11 or 12. Since 17-year-olds are enrolled in grades 9 through 12, this school population must be sampled in order to avoid statistical bias. It is expected that on the average these schools will contain five 17-year-olds. When a "special" school is selected for the sample all of its eligible students will be given an exercise package.

number of school groups or strata within each region that is possible if two schools are selected from each stratum. This latter condition allows for estimates of sampling variability to be computed from the assessment data.

D. Stratification Sample and Stratification of Pupils Within Selected Schools

The final step of the sample design will consist of stratifying the list of pupils supplied by each school into groups or strata so as to adequately spread or control the sample across grade and age level groups. For example, in assessing the 17-year-olds and eleventh grade populations, the following four strata would be constructed and randomly sampled:

1) 17-year-olds in 10th grade or less
2) 17-year-olds in 11th grade
3) 17-year-olds in 12th grade
4) 11th graders who are not 17-year-olds.

Strata 1, 2, and 3 would be sampled for 17-year-olds while strata 2 and 4 would be sampled for 11th graders. The allocation of the sample to the strata will be resolved after it is decided what statistical precision is expected for each of the two populations (17-year-olds and eleventh graders).

E. Summary of the Phase 1 Sample Design

Figure 4 graphically describes the sample design in its hierarchical or nested structure. The sample design consists of the following steps:*

Step 1 Group the schools by the ten geographical reporting regions.

Step 2 Split the schools of each reporting region into homogeneous groupings with respect to educational performance using the best data that is economically available. The number of groups to be formed is specified in Table 8.

Step 3 Select a random sample of two schools from each school group or strata formed in step 2 using a table of random numbers.

Step 4 Group the pupil list of each selected school into four strata defined by age and grade level.

Step 5 Select a random sample of pupils from each age group by grade level strata within the selected school using a table of random numbers.

---

* This sequence of steps assumed that the sampling frame or list of all schools, containing relevant stratification and sample selection data, has been constructed.

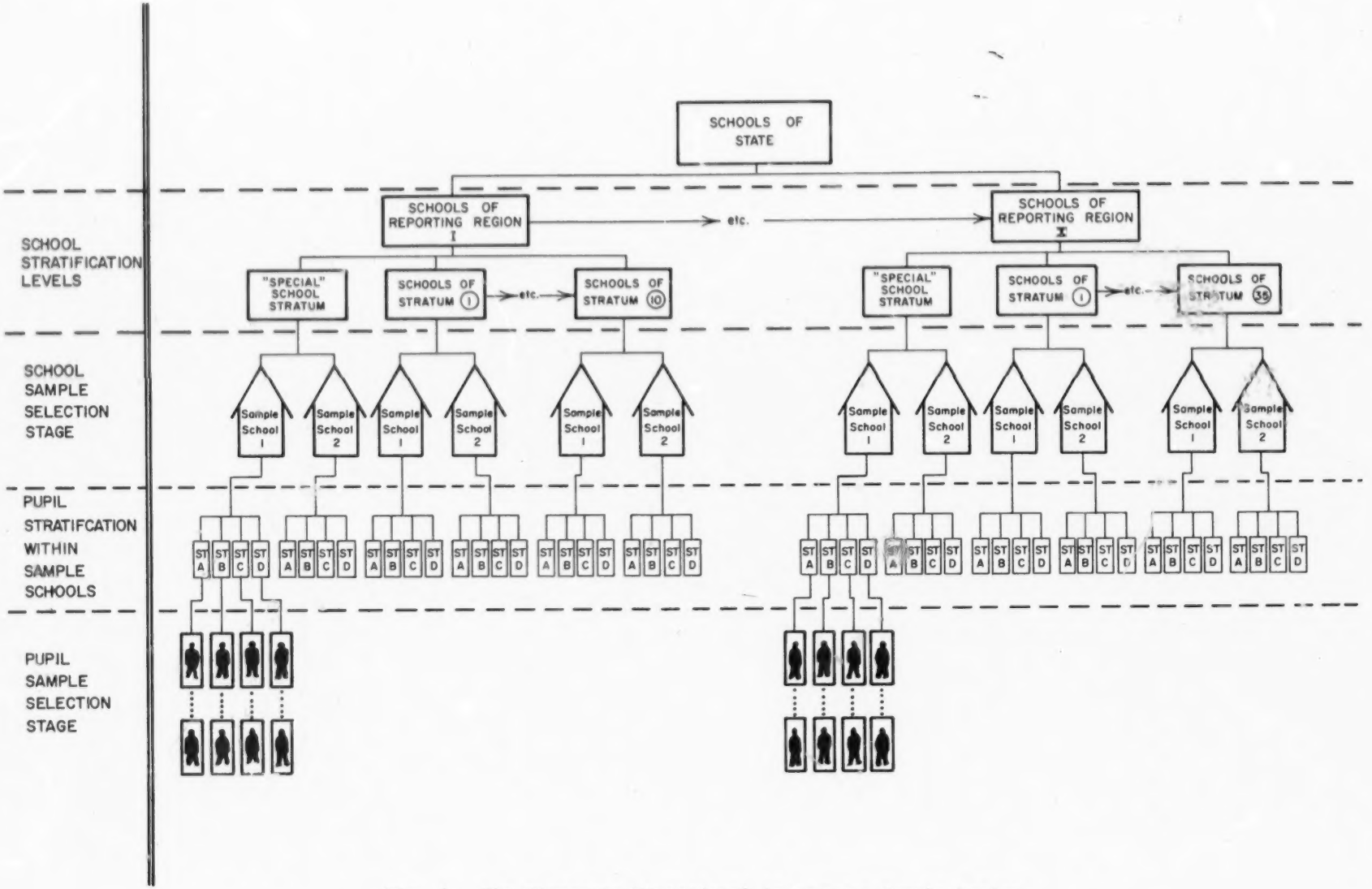


Fig. 4. The Minnesota Educational Assessment Sample Design

Chapter 6

Data Collection and Processing

I. INTRODUCTION

The quality of information gathered is greatly affected by how it is collected and who collects it. Since good decisions are rarely made on the basis of poor information, the task of collecting and processing data (includes editing and scoring) constitutes an integral aspect of the assessment plan.

Several questions were posed with respect to exploring alternatives for collecting and processing the Minnesota assessment data. Will state, school, or contracted personnel, or some combination of the three, serve as exercise administrators and supervisors? Are exercises to be administered to groups of students, individuals, or both? Will exercise administration instructions be taped? Will open-ended exercises be used? Will information pertinent to students, programs, teachers, and schools be obtained through actual site visits by survey teams, or through mail surveys? If information is to be gathered by survey teams, will these teams be comprised of state or contracted personnel, or some combination of the two? When will assessment data be collected? Will completed exercise booklets be optically scanned? What edit checks will be performed and how will errors be resolved? Will the assessment data be stored on magnetic tape files? Answers to these and similar questions were also evaluated with respect to resource constraints (time, funds, and personnel).

As a result, the approach to data collection and processing as outlined in this chapter is recommended as the most cost-effective method for collecting quality assessment data on elementary and secondary education in Minnesota. The responsibility for getting the exercise packages printed and for completing the data collection and initial editing of work tasks for this phase of the assessment program lies within the Survey Operations Section; the scoring and machine editing and data reduction tasks are the responsibility of the Scoring, Data Reduction, Sampling, Data Analysis, and Report Preparation Section. Since the Minnesota Department of Education does not have machine scoring and data reduction capabilities, these tasks will have to be completed either by the Information Systems Division of the Department of Administration or by an outside contractor. In either case, the organization which will do the actual scoring and transformation of data to computerized files should also set up the printing specifications. Efficiency of this operation would be increased if one organization could print the questionnaires and exercise booklets, as well as score and/or reduce their contents to computerized files.

This data collection and processing plan presupposes that preassessment conferences and workshops for representatives of participating schools and districts would be conducted by the Minnesota Department of Education and/or the contracted agency responsible for data collection. The primary purpose of these workshops will be to brief the representatives on the history, goals, purposes, and procedures of the assessment plan.

II. DATA COLLECTION

A. General

A field survey approach analogous to that used in the ongoing NAEP Program [Refs. 4 and 5] is recommended as the most cost-effective way for Minnesota to collect their assessment data. In this approach, the exercise packages are administered and the school questionnaires are collected by trained Exercise Administrators using paced tapes that contain all directions, exercises (except for the Reading exercises which are dependent upon Reading ability), and student questionnaire items. These tapes will be paced to allow the proper amount of response time for each exercise. All information on the paced tapes will also be printed in the exercise booklets and student questionnaires. Students will mark their responses directly into the packages.

These paced tapes help to assure the uniformity of administration by different exercise administrators and by the same exercise administrator from one administration to another. They also help to prevent exercises for subject areas other than Reading from measuring reading ability instead of the subject area being assessed.

The use of this field survey approach to collect data from a statewide sample of students also minimizes disruptions to the daily instructional routines of participating schools.

As previously indicated in Chapter 3 (Section II.B.), field survey operations will be conducted under the general supervision of the Head of the Survey Operations Section. Two District Supervisors and eight Exercise Administrators will be responsible for data collection. These personnel will form two teams, each consisting of a District Supervisor and four Exercise Administrators. A team will be assigned to a geographic area which would constitute approximately one-half of the work load for each six-week exercise administration cycle. The general job requirements for each of these positions were given in Chapter 3. Ensuing discussions as to how each survey team and each school in the sample would participate in data collection will provide a better understanding of the job skills required for these positions. The Head of the Survey Operations Section would be responsible for spelling out the entire data

collection process in a training manual, conducting training classes, and arranging work assignments and schedules for the District Supervisors.

The remainder of this section is directed specifically to the Phase 1 assessment; however, the operations for each subsequent fall, winter, and spring data collection cycle would be similar. The Phase 1 administration schedule requires that the Reading exercises be available in a form ready for printing by mid-February 1973 at the latest.

B. School Involvement

The field work would be scheduled and conducted so as to minimize school disruptions and demands imposed on school personnel. Every attempt would be made to collect as much data as possible from state and federal sources: i.e., those data routinely collected on reports, questionnaires, and surveys. Nevertheless, the following would be required of those schools selected in the Phase 1 assessment sample:

1) A roster of eligible* eleventh graders and of students whose birth dates lie on or between 1 October 1955 and 30 September 1956. (Similar requests would be made for each additional grade and age level assessed in subsequent assessments.) This information would be required in early February and would be used to draw the student sample.
2) One school official, preferably a counselor, to help arrange the testing schedule and to aid the District Supervisor and Exercise Administrators in having students at the administration site on schedule.
3) Adequate space for administering the exercises (exercise packages will be administered to an average of approximately 20 students from each school in the sample).
4) Cooperation of the principal in completing the school questionnaires (and of homeroom teachers in Phase 2 assessments to provide some of the background information for those elementary school students selected in the sample).

More specifically, the district superintendent of each school selected in the sample will be advised of their selection by mail in early February. This letter will furnish general information about the assessment and indicate that an assessment District Supervisor will make telephone contact with the superintendent within a short time. The purpose of the telephone contact will be to answer any questions and to arrange a meeting between officials of the individual schools, the superintendent, and the District Supervisor.

---

* Three types of students would be excluded from the target population even though they do meet the age and the grade level definitions: (1) non-English speaking students; (2) educable mentally retarded students; and (3) functionally disabled students. Principals will be provided with guidelines for identifying these students.

At this meeting, the school's involvement in the assessment program will be explained in detail. Space for administering the exercises and the aid of a school counselor will be arranged. Sampling procedures will be discussed and possible candidates for the Exercise Administrator positions will be solicited.

During the six week assessment period (15 March to 1 May), District Supervisors will supervise testing in several schools during any one week period. Assessment will generally be completed in one-half of a day for a single school--two hours of Reading exercises are planned for the Phase 1 assessment.

C. Package Administration

The collection of assessment data from students within each school will be done primarily by Exercise Administrators who have been hired and trained by the District Supervisors. District Supervisors, as time permits, will assist Exercise Administrators in performing this function. These Exercise Administrators will have been hired locally and should be familiar with the schools in their assigned areas. However, as the location of schools in the sample varies during subsequent assignments, judgemental decisions pertinent to hiring new Exercise Administrators or retaining the experienced ones will be required; i.e., does one pay more for travel expenses and retain experienced Exercise Administrators who would have to work outside of their localities when a new sample is drawn, or does one reduce travel expenses by hiring and training new Exercise Administrators who live in the proximity of the schools in the new sample? Individual decisions will have to be made on the basis of each Exercise Administrator's qualifications and travel schedule.

Each Exercise Administrator, after confirming the assessment schedule with school officials, would arrive at the school with a complete set of materials (school questionnaires, exercise booklets, paced tapes, and tape recorders), as well as note pads and pencils. Exercises would be administered to no more than 20-25 students. In the event that the number of assigned sample students exceeds this number, additional group sessions will be scheduled. The Exercise Administrator will also have specific instructions for scheduling make-up sessions if a minimum number of students are not available for testing.

The Exercise Administrator will leave the school questionnaire with, and explain it to, the principal on the same day that the exercises are scheduled for administration in his school. This will be done prior to the administration of exercises to the students. The completed school questionnaires are to be picked up from the principal on the same day, after the exercise administrations have been completed.

After the administration, the Exercise Administrator will (1) code certain information on the completed exercise packages, (2) collect the completed school

questionnaires from the principal, (3) conduct a brief field edit on all completed materials (it is always less expensive and more effective to immediately resolve errors or omissions on site), and (4) deliver all materials to the District Supervisor.

An important consideration in data collection--as well as in all other tasks associated with the state assessment--is the confidentiality of results. Utmost care must be taken by each person associated with the project to ensure that the confidentiality of the Minnesota assessment materials is protected at all times. Completed and uncompleted packages and other assessment materials are not to be given to anyone who is not actually involved in exercise administration. No duplication of materials is permitted. The names of the students are not to be associated with exercise packages; for example, space for the name of the student will not be provided in any part of any completed package. (Each exercise package will have the name of the student printed on the cover page. As each student turns in his completed exercise booklet, this cover sheet will be removed from the booklet.) Anything identifying students by name should not be removed from the school premises. Only those students assigned to the sample will participate in assessment. Other students will not be permitted to see the assessment packages.

III. DATA PROCESSING

A. General

Information will be gathered on machine readable/scorable forms in order to minimize errors inherent in reproducing or transferring data to magnetic tape files for computerized retrieval. However, editing and error resolution activities that involve more than one individual and require various levels of judgment will take place during several stages in data processing, both before and after the exercises have been scored. As mentioned earlier, it is important to safeguard the rights of the students and school principals involved in the assessment program. Therefore, certain measures are included in the data processing plan to ensure the confidentiality of information obtained on individual students and schools during the course of the assessment. These measures take into consideration the guidelines developed under the auspices of the Russell Sage Foundation for the collection, maintenance, and dissemination of pupil records. [Ref. 8]

B. Data Editing

The first editing check, as mentioned above in Section II, is a scan edit by the Exercise Administrator as the school questionnaire and exercise booklets are received at the testing site. This edit, done manually against a prepared instruction sheet,

should catch obvious omissions and errors. A "spot check" edit against the same instruction sheet will also be performed by the District Supervisor as he collects the booklets and questionnaires from each Exercise Administrator.

Following these two edits, the data would be delivered by the District Supervisor to the Department of Education for receipt control and final editing prior to being shipped to the scorer.

As these materials are logged in for receipt control at the Department of Education, they will be given a third scan edit for completeness. Follow-up activities would be initiated with schools for missing and/or incomplete materials and errors which cannot be reconciled by the central staff. All necessary coding on all questionnaires and booklets will be checked and, if necessary, altered to assure machine readability. After the return of all assessment materials has been satisfactorily completed, they will be packaged and shipped to the scorer for scoring and data reduction.

Before any items are actually scored, a brief pre-machine/-scoring edit should be performed on each exercise booklet by the scorer's staff. That is, the exercises and student and school questionnaires should be checked for illegal codes, inappropriate information, etc. Open-ended exercises should be scored by professional scorers who evaluate each student's answer and assign a proper numerical code indicating the appropriateness of the response to the exercise package. If any supplementary keypunching is required in processing, the open-ended types of data should be keyverified.

The exercise packages will then be machine scored and merged onto magnetic tape files with the optically scanned information from the school and student questionnaires. The scorer must establish and maintain certain quality control features in the optical scanning of the source documents. Several edit, or quality control checks, should be conducted during the scoring to ensure the required accuracy.

The magnetic tape files prepared by the scorer must be designed to provide for linking individual student data, including sample weighting information, to school data. Additional machine edit checks should be conducted on these files to ensure the accuracy, completeness, reasonableness, and consistency of the individual data files. Suspect files should be checked back against the original source document and resolution made at that time. The scorer is responsible for preparing clean tape files that include a complete record of the student's responses on the assessment instruments, as well as appropriate documentation of tape content and layout in accordance with approved documentation standards.

As a function of editing, a complete re-edit of the data files should be conducted at any stage in which corrections have been incorporated as a result of the checks and controls as stated above.

Note that the structure of the data in these files precludes anyone from associating a student name, or the name of the school that he attended, to an individual student data set. That is, each school data set and each data set for a student in that school will be coded by the same number. These numbers, though they are unique, cannot be related, by name, to a specific student or school.

C. Error Resolution

The resolution of errors identified in the editing process may take many forms, depending on the nature of the error and the editing stage in which it was noted; e.g., all errors noted by the Exercise Administrator at the test site should be corrected, whereas the time and expense involved in correcting an error noted at the scoring site might result in deleting the specific data item(s) in question. In addition, the scoring and developing of the files should be conducted so that those schools falling in the sample with an early closing date are scored and edited first. This allows for the resolution of any type of errors, particularly those of a critical nature, which must be resolved before the closing of schools.

The standard procedure for making all types of corrections would initially be the use of the source document (i.e., the exercise booklet and questionnaires) to determine what data should be appearing in the record. In addition to this initial check, the resolution of suspect data will generally take one or more of the following forms:

1) Follow-up action with data source and/or respondent.
2) Acceptance of data as reported.
3) Deletion of the specific data items in question or of the entire form.
4) Data imputation; i.e., the derivation of new data values based on specific rules and the values of related data items. Several imputation procedures such as cold-deck, hot-deck, and regression methods are available for use. However, the appropriateness of each of these methods will depend on the nature of missing data. In some situations involving probability sampling, it may be advisable to adjust the weights used instead of resorting to imputation. The imputation procedure to be used will depend on the nature of analyses, availability of auxiliary information available, and the pattern of missing data.
5) Confirmation of a computer generated correction.

Within these general guidelines, a specific set of error resolution rules should be established during the operational phase of the assessment for resolving errors in sets of specific data items. A file should be kept of all receipt control statistics and summary data on the detection and resolution of errors. This type of information

is extremely important to the refinement of data collection and processing procedures for subsequent assessments.

IV. SUMMARY

Data for the assessment program will be collected by specially trained survey teams using a "one day in--one day out" approach. These teams of District Supervisors and Exercise Administrators, under the supervision and coordination of the Head of the Survey Operations Section, will administer exercises and collect background information on the students and schools in the statewide sample. Data collection at each school will require no more than one-half a day and will be conducted with minimal disruption to school programs and slight impositions on students, teachers, principals, and other school officials. Exercise Administrators will use paced tapes to better ensure the standardization of all data collection and, except when assessing Reading, to help prevent the exercises from measuring reading ability instead of the subject area being assessed.

Assessment data will be collected on machine readable/scorable forms in order to minimize errors inherent in reproducing or transferring data to magnetic tape files for computerized retrieval. However, editing and error resolution activities that require various levels of judgment and involve more than one individual will be conducted at several phases in the data handling process.

Because it is important to safeguard the rights of participating students and school principals, certain measures are included in the data collection and processing plan to assure confidentiality with respect to all information collected on individual students and schools.

As such, these data collection and processing procedures assure a high degree of cooperation and a great degree of quality control with respect to both sample selection and data collection.

Chapter 7

Data Analysis Procedures

I. INTRODUCTION

In previous sections of this report, attention has been focused on the nature and methodology of the collection and processing of Minnesota educational assessment data. The role of these aspects of the program is to provide data of high integrity which accurately reflects the achievement status of Minnesota students. This chapter, with its concern for various strategies of statistical data analysis, focuses on the various procedures recommended to extract the most important and relevant descriptive measures from the data, as well as to detect important differences between various subgroups of students.

Three related discussions that are presented elsewhere in this report are important to this discussion of data analysis procedures: (a) the discussion of reporting variables in Section V of Chapter 2; (b) the desired outcomes discussion in Section VII of Chapter 2; and (c) the discussion of statistical precision in Sections II and III of Chapter 5. Reference will be made to these discussions in lieu of repeating them.

The data analysis procedures are presented in terms of three general sets of analyses. The first involves a descriptive analysis of the items of the student and school questionnaires. The second set of analyses involves comparisons of Minnesota Assessment results to NAEP results for the Nation and for the Central Region. The computation and comparison of assessment results for various reporting groups within the State of Minnesota constitutes the third basic set of analyses.

Tabular summaries of fictitious assessment data are also provided as examples of how the results of various analyses could be summarized. These tables are placed together at the end of the chapter.

II. PROCEDURES

A. General

The first step in any discussion of data analysis involves the classification of variables into types which are descriptive of their roles in the various statistical procedures which are to be used. The first type of variable is the dependent or output variable. These variables are the specific entities in terms of how statements about the results of the assessment will be formulated. For example, the response to each exercise by a respondent (1 for a "correct response," 0 otherwise) is a dependent variable which is used to compute p-values or proportions of the respondents within

each reporting group who gave an acceptable response to each exercise.*

The second type of variable is the independent variable in terms of which variations in the dependent variables are to be explained. The previously defined reporting variables (Chapter 2, Section V) serve as examples of independent variables.

The ensuing discussion will be rather general with respect to specific independent variables and combinations of independent variables to be used in the analysis since a final selection of reporting variables is still pending. In addition, the results of the preliminary analyses of the sample sizes in the various subgroups defined by the reporting variables will be important in selecting the various combinations of variables that can be used meaningfully as independent variables.

B. Descriptive Analysis of Pupil and School Background Items

The initial step in the analysis plan would be to describe the input characteristics of the students and schools in the state by summarizing the information collected through the student and school questionnaires. That is, the estimated proportions of Minnesota students in each grade/age level who are in each of the discrete categories corresponding to the reporting groups of the selected 20-25 reporting variables would be computed. Similarly, the estimated proportion of Minnesota schools in each reporting group for the school related reporting variables would be computed; i.e., the proportion of the state schools with various levels of per pupil expenditures, with various staff characteristics, with various levels of enrollments, with various student/teacher ratios, etc. If principals are asked to rate the adequacy of their facilities and staff with respect to certain attributes, the proportion of schools for each rating category would also be computed and reported.

Estimated proportions would also be computed for various combinations of reporting variables; e.g., the proportion of 11th grade students in each of three SES categories who attend schools in each of four different types of communities. The combinations of reporting groups that are feasible, from the viewpoint of statistical precision, will depend upon the size of the sample for each group (refer to Chapter 2, Section V.B.). A final selection of two-way and three-way tables to be presented should be based on their potential use in the explication of results and for their general informational value.

---

* More specifically, a p-value is an estimate of the total number of students in a group in terms of a population count who would have given an acceptable response, divided by the estimated total number of students in the group. Multiplied by 100, a p-value gives the percent-correct of the exercise. Notice that a p-value is not derived by dividing the number of respondents in the sample that give acceptable responses by the total sample size because all p-values are calculated using weighted responses based on the chance each respondent was selected into the sample.

C. Comparisons with National Assessment

Exercise p-values for various groups of Minnesota students would be contrasted with analogous groups of students in the Nation and in the Central Region. Such contrasts can be made only for those released NAEP exercises that are adopted by Minnesota. P-values for exercises that were administered by NAEP in individual sessions should only be compared with extreme caution to Minnesota p-values that were obtained by group administrations of the same exercises.

The sample is designed to provide Minnesota versus Nation and Minnesota versus Central Region comparisons for students at three age levels; i.e., 9, 13, and in-school 17-year-olds. In addition, the sample would be large enough to provide Minnesota versus NAEP comparisons for subgroups of 9, 13, and in-school 17-year-olds defined by sex, by four levels of parent's education, and by size of community. If desired outcomes measures are developed for Minnesota in the form of individual exercise p-values, they could also be compared with the Minnesota and NAEP results.

Table 9 is "mock-up Table" of fictitious Reading assessment results for 9-year-olds that serves as an example of how the Minnesota versus NAEP results versus desired outcomes might be displayed. (As discussed in Section VII of Chapter 2, it might not be feasible to collect desired outcomes data by age level.) The comparative results are presented by individual exercises, but since the exercise themselves are rather lengthy, the exercises and the percent responses for each foil would be placed in an appendix. Note in this table that NAEP comparisons are not possible for those exercises that would be developed by Minnesota to supplement the released NAEP exercises. Differences between the p-values of the state and the nation would also be analyzed for important statistical relationships and noted appropriately.

Table 10 provides an example of summarizing the Minnesota versus NAEP comparison by major themes. These fictitious results indicate the number of exercises within each theme for which the p-values for Minnesota students at each of the three age levels were equal to or greater (as judged by the standard error of their difference) than the results for the comparative NAEP groups, as well as for the state desired outcomes measures.

D. Within Minnesota Comparisons

These analyses would be concerned with computing and contrasting p-values for the various groups of students within Minnesota as defined by the reporting variables. Two levels of analysis would be used in making these comparisons. The first of these is similar to the analysis discussed in the previous section for making Minnesota versus NAEP contrasts, except that these comparisons will be made for the reporting groups within Minnesota. The first level of analysis is primarily descriptive in

character and involves the calculation of p-values and their respective standard errors for each exercise for the various reporting groups. In addition, analyses will be performed for evaluating the statistical relationship of certain comparisons across the groups of a reporting variable and between exercises within a reporting group. At the second level, relationships between the achievement measures (exercise responses) and student and school background variables would be investigated by analytic techniques such as Tukey's method of balanced fits. [Ref. 9] These analytical techniques involve the simultaneous examination of a number of independent variables and are useful for examining the effects of one independent variable while controlling or adjusting for the effects of other independent variables.

1. Level One Analysis

Weighted p-values and their standard errors will be computed for subgroups of students as defined by the reporting variables. Differences between the p-values of these subgroups will be analyzed.

Table 11 provides an example of reporting the results of the level one analysis for students in schools with different student enrollments. In addition to presenting p-values for each of the enrollment levels, the p-value for the state as a whole and the desired outcomes p-value for the state (if it is developed) are also included as reference points. Statistical differences between pairs of p-values would also be appropriately indicated on these tables. Standard errors would be reported in the appendix.

Results for the Minnesota reporting groups could also be summarized by major themes as per a format similar to that shown previously in Table 10.

It is also desirable to report the achievement status of student groups that have been simultaneously classified by any two or more of the previously discussed reporting variables; e.g., race by type of community or type of community by region, etc. Table 12 provides an example of exercise-by-exercise results for Race by Type of Community; Table 13 summarizes the results by Theme, in terms of the number of exercises on which the Minnesota goals were equalled or excelled. The results in Table 13 could also be summarized in terms of the median number of exercises correct under each major theme.

Tables such as Table 12 and 13 can be extremely useful in further pinpointing groups of students who may or may not have particular educational needs; e.g., Whites attending schools in Small Town/Rural communities may score much lower in "Word Attack Skills" than Whites in the other two Types of Communities. One might also use these tables to gain additional insight into possible interaction effects between two or more reporting variables.

2. Level Two Analysis

Whereas the level one analysis is directed at answering questions of the "what" dimension, the level two analysis is directed to the "why" dimension. That is, what is the extent of the relationships between achievement measures (exercise responses) and selected student, family, and school characteristics? Presentation of results in the form of two-way tables such as Tables 11 and 12 provides important insights into the interrelationships of dependent variables to underlying independent variables. However, multiple regression and certain multivariate techniques are more powerful tools for unraveling the complex relationships among a set of independent and dependent variables.

The purpose of the level two analysis is to attempt to glean information from the assessment that is not revealed in one-way or two-way descriptive tables. Analytical techniques such as balanced fits attempt to supply answers to such questions as: what would be the differences in p-values between students in schools with large enrollment differences if the distribution of students by parental education, race, SES, type of community in which school is located, etc., had been the same for all schools? A great danger in this approach is attributing causality to those variables that turn out to be significant in the statistical model. Nothing short of experimentation, if that, can demonstrate what the actual effects of these variables would be. Statistical models based on survey data (non-experimental data) can only provide hints and insights into the probable effect of these variables on achievement. Nevertheless, it is important to look at the data in every way for hints about how the education system works and how it can be improved. The level two analysis involves statistical tools for doing this.

This level of the analysis involves the formulation of prediction equations for both qualitative and quantitative dependent variables as a function of various independent variables. At this stage, one would be trying to determine how a subset of independent variables work together in producing variation in achievement. For example, the unique contribution of student background and school factors in predicting achievement can be isolated. In other words the multiple correlation coefficient for the prediction equation can be decomposed into components which are unique to specified predictor variates, and those which are shared by various combinations of predictor variates. This is the approach taken in the recent reanalysis of the Coleman data. Criterion scaling can also be used to linearize the relationship between various independent and dependent variables. The ultimate objective of these complex analyses is the derivation of prediction equations from the sample data that are unbiased and efficient estimates of the relationships for all Minnesota students in a particular grade or age class.

The results of the level two analysis will be displayed in tables similar to those of Tables 11 and 12, except that the p-values will have been adjusted for differences in selected school, student, and process variables other than those used to define the particular reporting groups of that table. For example, the results in Table 11 might be adjusted for variations in per pupil expenditures, student and family background characteristics, and type and size of the community variables that might exist between the students in each of the five enrollment categories.

III. SUMMARY

The data analysis plan for the proposed Minnesota Assessment program consists of three general sets of analyses.

The first is a descriptive analysis of the questionnaire responses that would serve to describe the input characteristics of the students and schools. That is, estimated proportions of Minnesota students in each grade/age level who are in each of the discrete categories of the selected reporting variables would be computed. Similar estimates would also be computed for the proportion of Minnesota schools in each reporting group for the school related variables. In addition, estimated proportions would also be compiled for various combinations of reporting variables.

The second set of analysis involves comparisons of exercise p-values for various groups of Minnesota students to those of students in comparable age groups (9, 13, and in-school 17-year-olds) from throughout the Nation and Central Region who participated in NAEP.

The third set of analysis is the contrasting of exercise p-values for the various groups of students within Minnesota as defined by the reporting variables. Two levels of analysis would be conducted in this phase of the analysis. The first level is primarily descriptive in orientation and would involve the calculation of exercise p-values, their respective standard errors, and statistical comparisons between certain pairs of p-values. At the second level of analysis, relationships between achievement measures and selected student and school background variables would be investigated.

Table 9

CORRECT RESPONSE RATE (P-VALUES) BY READING EXERCISES FOR 9-YEAR-OLDS IN MINNESOTA, THE NATION, AND CENTRAL REGION
(Fictitious Results Prepared for Display Purposes Only)

Reading Skill Area[1]	Minn. Desired Outcomes (% Correct)	Percent Correct Responses		
		Minn.	Nation	Central Region
I. Word Attack Skills				
Theme 1, Initial Consonants				
Item 1	80	84	70	74
Item 2	75	77	68	62
Item 3	85	82	NA[2]	NA[2]
.				
.				
.				
Theme 2, Syllabication				
Item 1	85	80	67	81
Item 2	90	93	88	77
.				
.				
.				
II. Word and Sentence Comprehension				
Theme 3, Word Meanings				
Item 1	90	73	NA[2]	NA[2]
Item 2	80	70	68	78
.				
..				
.				

[1] Specific exercises would be presented in an Appendix.

[2] These exercises were developed by Minnesota and were not in the packages administered by the National Assessment of Educational Progress Project.

Table 10

MINNESOTA ASSESSMENT RESULTS VERSUS NAEP NATIONAL AND CENTRAL REGION BY READING SKILL AREAS AND AGE LEVELS
(Fictitious Results Prepared for Display Purposes Only)

Reading Skill Area	Number of Items in Which Percent Correct Responses for Minnesota Equaled or Exceeded (1) State Desired Outcomes, (2) National NAEP Results and (3) Central Region NAEP Results: 9-Year-Olds			13-Year-Olds			In-School 17-Year-Olds		
	Minn. Outcomes	Nation	Central Region	Minn. Outcomes	Nation	Central Region	Minn. Outcomes	Nation	Central Region
I. Word Attack Skills	9 of 20	7 of 9	5 of 9	21 of 30	10 of 13	7 of 13	28 of 40	11 of 14	8 of 14
Theme 1, Initial Consonants	5 of 12	3 of 5	2 of 5	10 of 14	4 of 6	3 of 6	13 of 20	5 of 6	4 of 6
Theme 2, Syllabrication	4 of 8	4 of 4	3 of 4	11 of 16	6 of 7	4 of 7	15 of 20	6 of 8	4 of 8
II. Word and Sentence Comprehension	11 of 15	10 of 13	7 of 13	14 of 20	9 of 12	7 of 12	20 of 27	12 of 15	9 of 15
Theme 3, Word Meanings	11 of 15	10 of 13	7 of 13	14 of 20	9 of 12	7 of 12	20 of 27	12 of 15	9 of 15
III. Reading Study Skills	19 of 30	9 of 14	6 of 14	29 of 45	11 of 15	9 of 15	41 of 60	21 of 28	14 of 28
Theme 4, Visual Aids	8 of 11	5 of 5	4 of 5	14 of 18	7 of 9	6 of 9	16 of 22	8 of 10	7 of 10
Theme 5, Written Directions	4 of 9	1 of 4	1 of 4	6 of 14	2 of 4	2 of 4	12 of 20	8 of 9	4 of 9
Theme 6, Reference Materials	7 of 10	3 of 5	1 of 5	9 of 13	2 of 2	1 of 2	13 of 18	5 of 9	3 of 9
IV. Comprehension in Longer Discourse	19 of 40	15 of 25	9 of 25	25 of 50	19 of 28	14 of 28	37 of 60	20 of 26	13 of 26
Theme 7, Reading for Facts	5 of 9	4 of 5	2 of 5	6 of 12	4 of 7	3 of 7	9 of 15	6 of 6	3 of 6
Theme 8, Reading for Main Idea and Organization	8 of 12	7 of 9	4 of 9	8 of 14	6 of 8	4 of 8	13 of 18	6 of 8	4 of 8
Theme 9, Reading and Drawing Inferences	3 of 11	2 of 6	2 of 6	7 of 14	4 of 6	3 of 6	6 of 14	6 of 7	5 of 7
Theme 10, Critical Readings	3 of 8	2 of 5	1 of 5	4 of 10	5 of 7	4 of 7	9 of 13	2 of 5	1 of 5
Total	58 of 105	41 of 61	27 of 61	89 of 145	49 of 68	37 of 68	126 of 187	64 of 83	44 of 83

Table 11

CORRECT RESPONSE RATE (P-VALUES) BY READING EXERCISES AND SCHOOL SIZE FOR FOURTH GRADERS IN MINNESOTA
(Fictitious Results Prepared for Display Purposes Only)

Reading Skill Area[1/]	Minn. Desired Outcomes (% Correct)	Percent Correct Responses — School Enrollments					
		< 200	200-499	500-799	800-1097	1100 +	Minn.
I. Word Attack Skills							
Theme 1, Initial Consonants							
Item 1	80	84	70	72	74	78	73
Item 2	75	77	68	66	62	65	64
Item 3	85	82	73	77	76	74	72
.							
.							
.							
Theme 2, Syllabication							
Item 1	85	80	67	69	71	74	70
Item 2	90	93	88	89	77	63	59
.							
.							
.							
II. Word and Sentence Comprehension							
Theme 3, Word Meanings							
Item 1	90	73	70	83	84	79	76
Item 2	80	70	68	72	78	59	65
.							
.							
.							

1/ Specific exercises would be presented in an Appendix.

Table 12

READING RESULTS (PERCENT CORRECT RESPONSES) IN WORD ATTACK SKILLS FOR MINNESOTA 17-YEAR-OLDS BY RACE AND TYPE OF COMMUNITY
(Fictitious Results Prepared for Display Purposes Only)

Reading Word Attack Skills Theme	Exercise	Race	Type of Community		
			Large City	Large City Fringe/ Medium City	Small Town/ Rural
1. Initial Consonants	1	Black	50	60	45
		White	65	70	60
		Indian	55	58	52
	2	Black	40	65	41
		White	73	82	58
		Indian	65	67	45
	. . .				
2. Syllabication	1	Black	45	49	39
	2	White	70	78	48
		Indian	49	53	41
	. . .				
. . .	. . .				

1/ Specific exercises would be listed by major theme in an Appendix of the Report.

Table 13

READING RESULTS (PERCENT CORRECT RESPONSES) FOR MINNESOTA 17-YEAR-OLDS BY THEME, RACE, AND TYPE OF COMMUNITY (TOC)

Basic Reading Skill Area		Number of Items	Race	Number of Items in Which TOC Percent Correct Responses By TOC Equaled or Exceeded Minnesota Expectancy		
Area[1]	Theme			Large City	Large City Fringe/Medium City	Small Town/Rural
I	1. Initial Consonants	20	Black	8	11	3
			White	14	15	8
			Indian	10	12	4
	2. Syllabication	20	Black	9	11	7
			White	15	18	10
			Indian	13	12	6
II	3. Word Meanings	27	Black	8	10	5
			White	13	18	8
			Indian	11	12	6
III	4. Visual Aids	22	Black	11	12	8
			White	14	17	12
			Indian	12	13	10
	5. Written Directions	20	Black	7	8	6
			White	12	15	10
			Indian	9	10	6
	6. Reference Materials	18	Black	6	8	4
			White	9	10	6
			Indian	8	8	5

-Continued-

Table 13 (Continued)

Basic Reading Skill Area[1]	Theme	Number of Items	Race	Number of Items in Which TOC Percent Correct Responses By TOC Equaled or Exceeded Minnesota Expectancy		
				Large City	Large City Fringe/Medium City	Small Town/Rural
IV	7. Reading for Facts	15	Black	6	7	4
			White	10	12	6
			Indian	8	9	5
	8. Reading for Main Idea and Organization	18	Black	8	9	6
			White	12	13	9
			Indian	10	11	7
	9. Reading and Drawing Inferences	14	Black	7	6	3
			White	9	10	5
			Indian	8	8	4
	10. Critical Readings	13	Black	4	6	1
			White	8	10	3
			Indian	5	8	2

[1] The Reading Basic Skill Areas are defined as follows:

I. Word Attack Skills
II. Word and Sentence Comprehension
III. Comprehension in Longer Discourse
IV. Reading Study Skills

Chapter 8

Reporting and Dissemination

I. INTRODUCTION

The basic purpose of a dissemination plan for statewide assessment results is to insure that accurate information is made available to all interested people in the state, at a level of sophistication (detail) commensurate with their background, needs, and purposes. This information is needed so that people at all levels can evaluate properly the need for change, so that they have enough information of the right kind and type to make intelligent and data-based educational decisions. Thus a legislator or high-level policymaker would have a far different background, need, and purpose for assessment results than would a housewife with children in school. Too little information in the former case would be deleterious while too much in the latter case would provide unneeded, unwanted, and perhaps misunderstood data. This means, then, that report formats, audio-visual aids, presentation modes, and any other form of promulgation and publication decided upon and used should be tailored to some degree to decisionmaking groups, special interest groups, and the general public. The final report of the Minnesota Educational Assessment Program for each year or cycle may be a somewhat bulky, detailed, statistical narrative which should be interpreted for a wide variety of potential audiences.

There are at least four levels which should be considered in understanding the various needs of the consumers of assessment program results. The policymaking level consists of legislators and members of the executive branch of the state government who are charged with the responsibility for establishing broad policies which, when carried out, will best meet the educational needs of the state. For example, policymakers may decide on the basis of statewide results that a major emphasis should be put on Citizenship rather than some other area which has been receiving emphasis but which exceeds expected levels.

The decisionmaking level is comprised of personnel throughout the state in the Department of Education who are charged with carrying out policy through making the basic operational decisions and allocations of resources. These are the various District Superintendents, curriculum directors at the state level, curriculum specialists at various levels, and principals of schools. In the Citizenship example, decisionmakers would be concerned with finding and recruiting appropriate specialists in Civics and other social sciences, developing goals and objectives for new areas of instruction, and devising curriculums to meet these new objectives. The assessment results would

provide these decisionmakers with some insight into specific areas on which Citizenship emphasis should be brought to bear.

Third is the operational level, the "on the street" level in which educators work actively with children and parents, face the very real problems in modern education, and carry out all the policies and decisions which have been made. These are teachers, school-based curriculum and guidance specialists, and other teaching staff. People at this level are devising and carrying out lesson plans and classroom projects, and bringing outside experiences unique and meaningful to their own groups of students. The broad range of Citizenship needs revealed in the statewide assessment, for example, would have to be translated into concrete behavioral terms both for this level and by this level.

Finally, the fourth level--the public level--brings us back full circle to the policymaking level. Consisting of local school board members, various special interest groups, both concerned and somewhat-less-than-concerned parents of children in school, and a very large group of adults with no children in school, this group through the democratic processes can and should have a strong effect on the first level. People at this level are potentially most helpful in deciding what is important for children to know, and what is mere trivia. Their perceptions will often differ markedly from those of subject-matter people at both the decisionmaking and operational levels.

The artifactual division or categorization into levels does not mean, however, that certain levels can have access to only certain bits of information interpreted in certain ways. Rather it means that disseminators must be alert to their needs and be ready to respond with appropriate detail when queried.

Under the alternative selected by Minnesota of three grade/age levels, two subject-matter areas, and ten reporting regions for the assessment program, the spring 1973 (Phase 1) assessment can be considered as a gearing up effort, inasmuch as only one area and age group--Reading for 17-year-olds--is being assessed. Subsequent years and cycles (Phase 2) will have the full scope of three grade/age levels and two subject matter areas. Hence, for dissemination purposes, spring 1973 should also be viewed as developmental in that the strategies tried for effective dissemination could be changed (if required) for subsequent assessments.

As a final comment, there is a difference in the desired impact of assessment dissemination and publicity in the beginning as compared to the end of a year or a cycle. The assessment program must be "sold" in different ways. Hence, the primary vehicle for wide dissemination and publicity is a series of Fall Workshops that consist of a results dissemination phase for just-completed administrations, and a planning and publicity phase for to-be-completed administrations. The exception to this will be Phase 1 which will have only a pre-assessment workshop.

II. PRE-ASSESSMENT WORKSHOPS FOR PHASE 1 (SPRING 1973)

As soon as possible after the formal approval of the assessment program, and before the actual administration of the Phase 1 assessment exercises in the schools, a series of 6-10 workshops should be conducted across the state for personnel at the decisionmaking level to insure that the goals, aims, and purposes of the program are fully understood. Reporting variables and reporting groups will be explained, as well as the basic elements of the Reading assessment program to be conducted for 17-year-olds. These pre-assessment workshops should include a slide and tape presentation so that consistent information would be promulgated. Appropriate officials from the Department of Education first should be trained on the overall assessment program, and then be available in these workshops to answer questions and expand on themes as necessary. Relevant assessment handouts should be distributed, and suggestions should be given for attendees to hold similar workshop discussions with teachers, specialists, and other personnel potentially involved with the program. Press releases should be generated so that the public would be informed.

The Department of Education should give serious consideration to holding similar informative workshops for the benefit of lay educational and community leaders, school board members, and the general public. A suggested method for doing this is through the use of a half-hour or less educational television program on tape which is repeated at intervals on ETV stations. In addition, an audio tape-and-slide show could be put together from the same basic materials and made available at no charge to such groups as local service organizations (Rotary, Kiwanis, Exchange, etc.), school boards, and relevant action or special interest groups.

III. PHASE 1 (SPRING 1973) DISSEMINATION

The basic document containing the results of the spring 1973 assessment in Reading for 17-year-olds will be the Technical Report, a statistical narrative report produced by the efforts of all Sections of the Assessment Office, but especially the data analysis and report preparation section. The Technical Report will present in analytical detail all aspects of the methodology of the assessment, such as instrument development processes and specifications; sample design stratification and weighting criteria; details of administration and scoring; statistics relating to nonresponse and sampling errors, and estimation and imputation procedures; and a discussion of the limitations of the data. In addition, there will be extensive tabulations of results, interpretations and analyses of these tabulations, and sufficient technical detail throughout to meet the needs of technically oriented readers.

Based upon the preliminary draft of the Technical Report, which should be available three to four months after exercise administration has been completed, a Highlights Report in popular language and format can be drafted. This Highlights Report would be widely disseminated to all schools which participated in the program, appropriate curriculum specialists in reading and language arts, district superintendent offices, state level Department of Education personnel, etc. Included in the Highlights Report will be an order blank for obtaining the Technical Report. This will assure wide availability of the Technical Report on request.

The ability of any and all persons to obtain the basic Technical Report is crucial. Misinterpretation of assessment results can occur on the basis of such necessarily condensed information as the Highlights Report will contain. The Technical Report will be initially distributed to such personnel and agencies manifestly having the need and ability to use and profit by it, such as the Commissioner, superintendents, legislative committee members, certain research and evaluation specialists on the Department of Education staff, and AAC and TAC members.

It is vitally important that assessment results, in order to have any substantial effect on the educational process, be interpreted in operational terms by and with operational decisionmakers. Therefore, it is proposed that the primary vehicle for interpretive dissemination be a series of Fall Workshops analagous to the pre-assessment workshops.

These workshops will be primarily geared to the needs and expectations of operational people in education; hence, assessment results should be interpreted on not only a local basis, but also in classroom and subject matter terms. For example in Citizenship a goal of understanding the rights and freedoms of individuals is by itself too broad to implement. Participants in workshops would have to interpret assessment results on this goal by generating new ideas to teach about, facts to impart, and attitudes to look for. Specifically in Phase 1, Reading assessment results will be worked over with reading, language arts, grammar, remedial reading, and composition instructors, to name but a few.

These Fall Workshops will be held either during the pre-school sessions of teacher and administrator preparation before the beginning of each school year, or early in the fall immediately before the first field administration. The first such workshop will be based on the spring 1973 (Phase 1) Reading results for 17-year-olds; and the audio-visual presentations (slides, tapes, handouts) would present not only these results in terms of the data, but possible interpretations of the data pointing to areas of strength and areas of weakness. For example, the results may show clear regional differences in reading level for which there appear no valid reasons, and which require prompt attention. Conflicting results may emerge which need clarification,

e.g., word attack skills may be low in an area where there is no accompanying low comprehension level. Concepts such as these must be clarified and joint decisions made on local actions to be taken where necessary.

Nearly equal in importance to dissemination of Phase 1 Reading results at these workshops is the presentation of lead-in material for the subsequent assessments of Phase 2. This part of the workshop would be similar in purpose to the workshops held previously in Phase 1, but geared specifically to the assessment administrations to be conducted next. For example, the Phase 1 dissemination workshop would also introduce the first year of the Phase 2 assessment, i.e., Reading in two grade/age levels and Literature in three grade/age levels.

About 6 to 10 such workshops again are envisioned, with perhaps 40 to 50 attendees each, spread over the 10 reporting districts of the state to minimize travel costs and time. Appropriate press releases should be created to publicize the workshops, and the media would be invited to attend.

IV. PHASE 2 DISSEMINATION

Throughout the cyclical schedule of the assessment program, there should be an effort to keep the program in the public eye through a series of press releases as milestones are passed. It should also be kept in the forefront of the education profession's vision. The major method for this will be the annual Fall Workshop presenting the previous year's results and the coming year's assessment program and schedule, together with a presentation of ideas and innovations arising out of the information gained through the assessment program.

Each assessment effort, e.g., Music and Science for 9-year-olds in 1975-76, or Citizenship and Art in grade 8 in 1977-78, will have both a Technical Report and a Highlights Report, the latter to have wide dissemination and the former to have both a limited initial dissemination and a wide availability on request. The workshop each year in the fall will serve to tie the annual, three grade/age level assessment efforts together and point to the coming year's efforts.

Finally, provisions will have to be made for a formal presentation to the Legislature and/or the Executive section of the government of the State, and other policymaking persons, near the end of each fiscal year so as to justify both the current assessment program expenditures and projected program expenditures. This presentation, though quite similar to the workshop presentations, would include more data relevant to the types of decisions these bodies find it necessary to make.

V. SUMMARY

A dissemination plan must be responsive to the informational needs of various levels of consumers of educational assessment results and of educational decisionmakers who will use assessment information. These consumers are individuals at the policy-making level, the decisionmaking level, the operational level, and the public level. All levels must have access to all results; but their specific information needs vary. The Technical Report and the Highlights Report will meet some of these different needs.

The primary vehicle for assessment dissemination and planning is seen as a series of Fall Workshops, each of which has a results dissemination aspect and a planning aspect. These two aspects will be split, however, for Phase 1.

A series of pre-assessment workshops is proposed in spring 1973, before actual Phase 1 assessment administration, to acquaint the decisionmaking and operational levels with the project. In addition, wide public exposure to the assessment project is suggested through television, service clubs, and special interest group presentations.

Actual dissemination of Phase 1 results will occur by means of the basic Technical Report, a widely distributed Highlights Report, and Fall Workshops for decisionmaking and operational persons in the subject area(s) assessed. These workshops will also serve a planning function for the early Phase 2 assessments.

Yearly Phase 2 dissemination will be in the same manner: Technical Report, Highlights Report, and dissemination and planning Fall Workshops on subject-matter areas. Each year there will also be special presentations prepared for the policy-making level.

REFERENCES

1. John W. Adams. Minnesota Educational Assessment Annual Report, 1971-72. Unpublished document available from Minnesota Department of Education, Division of Planning and Development, August 1972.

2. Minnesota State Department of Education. Minnesota Statewide Educational Assessment: Pilot Phase Results. St. Paul, Minnesota: Minnesota State Department of Education, Division of Planning and Development (in process of publication).

3. "Scotching a Dangerous Myth: Higher Expenditures, Poorer Results," Phi Delta Kappan, LIV, No. 1 (September 1972), p. 74.

4. Questions and Answers About the National Assessment of Educational Progress. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, April 1972.

5. Frank B. Womer. What is National Assessment? Denver, Colorado: National Assessment of Educational Progress, Education Commission of the States, 1970.

6. Carmen J. Finley and Frances S. Berdie. The National Assessment Approach to Exercise Development. Denver, Colorado: National Assessment of Educational Progress, Education Commission of the States, 1970.

7. Ralph W. Tyler. What Is An Ideal Assessment Program? Sacramento, Calif.: State Department of Education, Bureau of References Services, July 29, 1968 (Rev.), p. 9.

8. Guidelines for the Collecting, Maintenance, and Dissemination of Pupil Records. A Report on the Conferences convened by Russell Sage Foundation, Sterling Forest Conference Center, Sterling Forest, N. Y., May 25-28, 1969. Hartford, Connecticut: Connecticut Printers, Inc., 1970.

9. Reading and Literature: General Information Yearbook. Denver, Colorado: National Assessment of Educational Progress, Education Commission of the State, May 1972, pp. 83-85.

Appendix A

Membership Lists for the Minnesota Assessment Advisory Council and Technical Advisory Committee

Assessment Advisory Council Membership

Professional Educational Organizations

Minnesota Education Association - Mr. William M. (Mort) Mondale

Minnesota Federation of Teachers - Mr. Robert Knock

Minnesota Association for Supervision and Curriculum Development - Dr. Richard Kimpston

Minnesota Reading Association - Dr. Robert Schreiner

Minnesota Council of Teachers of Mathematics - Dr. Randall Johnson

Minnesota School Counselors Association - Dr. Jerry Thompson

Minnesota Elementary School Principals Association - Dr. Werner Tismer

Minnesota Association of Secondary School Principals - Mr. Ernie Gustafson

Minnesota Association of School Administrators - Dr. John B. Davis, Jr., and Dr. Daniel B. Mjolsness

Governor's Office

Mrs. Wenda Moore

Statewide Testing Service

Dr. E. Gary Joselyn

State Planning Agency

Mr. Dean Honetschlager

State Legislature

Senator Paul Overgaard, Vice-Chairman - Senate Committee on Education

Representative Harvey Sathre, Chairman - House Committee on Education

Representative George F. Humphrey, Vice-Chairman - House Committee on Education

Higher Education

Dr. Jack D. Merwin, Dean
College of Education
University of Minnesota

Non-Public Education

Mr. Robert D. Burke
Director of Research
Archdiocese of St. Paul-Minneapolis

Citizen Groups

Minnesota School Boards Association - Dr. Rollin Dennistoun

Minnesota Citizen's Committee on Public Education - Mrs. Allen Sulerud

Minnesota Congress of Parents and Teacher, Inc. - Mrs. E. E. Jacobsen

Minnesota League of Women Voters - Mrs. William Jones

Minnesota Association of Commerce and Industry - Mr. Richard A. Bragg

Hill Family Foundation - Mr. Robert Bonine

Minnesota-Dakotas State Conference of the National Association for the Advancement of Colored People - Mr. Curtis Chivers

Technical Advisory Committee Membership

Dr. Donald Sension
Coordinator for Research
and Evaluation
Hopkins Public Schools
1001 State Highway 7
Hopkins, Minnesota 55343

Dr. Roger Giroux
Director of Research
and Evaluation
Duluth Public Schools
226 North 1st Avenue, E.
Duluth, Minnesota 55802

Dr. Andrew Ahlgren
Associate Director
Center for Educational Development
317 Walter Library
University of Minnesota
Minneapolis, Minnesota 55455

Dr. Glen Bracht
Assistant Professor
Psychological Foundations
University of Minnesota
323 Burton Hall
Minneapolis, Minnesota 55455

Dr. Jack C. Merwin, Dean
College of Education
University of Minnesota
104 Burton Hall
Minneapolis, Minnesota 55455